

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:48:34 ; Search time 20 Seconds  
(without alignments)  
1211.725 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MAELEFVQIIIIIVVMVMV.....PLESAAIWSKDKQKGHPL 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125.5	9.3	634	2 T00388	hypothetical prote
2	123.5	9.2	508	2 T09046	proline-rich prote
3	122	9.1	530	2 T48627	hypothetical prote
4	113	8.4	628	2 S19150	hypothetical prote
5	111.5	8.3	1184	2 G01763	atrophin-1 - human
6	111.5	8.3	1184	2 S08332	atrophin-1 - human
7	111	8.2	727	2 C84534	hypothetical prote
8	110.5	8.2	778	2 I38487	tastin - human
9	109.5	8.1	1099	2 A56155	tumor suppressor p
10	107	7.9	2357	2 A59249	class VII unconven
11	105	7.8	1145	2 T33606	hypothetical prote
12	103	7.6	628	2 S01955	hypothetical prote
13	102.5	7.6	1357	2 T23265	hypothetical prote
14	101	7.5	384	2 S51796	vasodilator-stimul
15	101	7.5	944	2 T28734	hypothetical prote
16	101	7.5	1571	2 T00062	hypothetical prote
17	100.5	7.5	760	2 T06291	extensin homolog 1
18	100	7.4	363	2 T16755	hypothetical prote
19	100	7.4	445	2 S00256	Krox-20 protein -
20	100	7.4	628	2 JQ0110	hypothetical 69K p
21	100	7.4	1132	2 T49403	related to protein
22	99.5	7.4	786	2 T01456	extensin homolog F
23	99.5	7.4	872	2 S33015	hypothetical prote
24	99	7.3	448	2 T06076	proline-rich prote
25	99	7.3	571	2 T43456	hypothetical prote
26	99	7.3	1051	2 T51904	hypothetical prote
27	98.5	7.3	1366	2 B84924	hypothetical prote
28	98	7.3	761	2 C84672	hypothetical prote
29	98	7.3	1172	2 T00065	hypothetical prote

30	97	7.2	527	2 B70700	hypothetical prote
31	96	7.1	436	2 T15331	hypothetical prote
32	96	7.1	620	2 S06733	hydroxyproline-ric
33	96	7.1	710	2 D96728	hypothetical prote
34	96	7.1	895	2 C86371	99.7K hypothetical
35	96	7.1	1460	1 ED8E1F	immediate-early pr
36	96	7.1	1844	1 RHWPTM	genome polyprotein
37	95.5	7.1	707	2 A46302	PIB-associated spl
38	95.5	7.1	998	2 T30930	hypothetical prote
39	95.5	7.1	1240	2 JC5209	insulin receptor s
40	95.5	7.1	5762	2 A41819	proline-rich pepti
41	95	7.0	876	2 A49508	protein-tyrosine k
42	95	7.0	896	1 A35782	cytokine receptor
43	95	7.0	951	2 T47617	extensin-like prot
44	94.5	7.0	310	1 PIHUSD	salivary proline-r
45	94.5	7.0	322	2 S25299	extensin precursor

ALIGNMENTS

RESULT 1

T00388  
hypothetical protein KIAA0616 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T00388  
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.;  
DNA Res. 5, 169-176, 1998  
A>Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
A:Reference number: Z14142, MIM:29403880; PMID:9734811  
A:Accession: T00388  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-634 <ISH>  
A:Cross-references: EMBL:AB014516; NID:G3327045; PIDN:BAA31591.1; PID:G3327046  
A:Experimental source: brain  
C:Genetics:  
A:Gene: KIAA0616

Query Match	9.3%	Score 125.5;	DB 2;	Length 634;
Best Local Similarity	24.7%	Pred. No. 0.026;		
Matches	53;	Conservative	24;	Mismatches 63; Indels 75; Gaps 8;
QY	38	ISRHSQGRRE-----	DALSSGCLWPSESTVSGNGIPQVYAPRP	80
Db	314	LSLSTEARQQASPTLSPLSPTQAVAMDALSLEQLPYAFPTQAGSQPPPPPPPP	373	
QY	81	TDRLAVPPAQRERHFRQTPYVLOHEIDL-----	PTISLSDGEEPPP--YQGPCVL	132
Db	374	-----PPASQO-----	PPPPPPQAFVRLPGGPLLPSASLTRGPQPPPLAVTPSSL	421
QY	133	QLRDEQ-----	QLELNRESVAPPNRTIFDSDLMDSARLGGCPFPSSNSG	178
Db	422	QSPSPENPGQSMGIDIASAPALQYRTSAGSPANQS-----	PTSPVSNQG	467
QY	179	ISATCYGSGRMGPPTTYSEVIGHYPGSSFQHQ	213	
Db	468	FS-----	PGSSPQHTSTLGSVFGDAYVEQ	492

RESULT 2

T09046  
proline-rich protein F26K10.180 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 22-Oct-1999  
C:Accession: T09046  
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16533  
A:Accession: T09046  
A:Molecule type: DNA  
A:Residues: 1-508 <BEV>

A; Cross-references: EMBL:AL049803; GSPDB:GN00062; ATSP:F26K10.180  
A; Experimental source: cultivar Columbia; BAC clone F26K10  
C; Genetics:  
A; Gene: ATSP:F26K10.180.  
A; Map position: 4  
A; Introns: 52/3; 160/3

RESULT 3  
T48627  
hypothetical protein T15N1.30 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48627  
R:Sevan, M.; Murphy, G.; Bidley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24493  
A:Accession: T48627  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-530 <BEV>  
A:Cross-references: EMBL:AL163792  
A:Experimental source: cultivar Columbia; BAC clone T15N1  
C:Genetics:  
A:Map position: 5  
A:Introns: 83/3; 176/3  
A:Note: T15N1.30

RESULT 4  
S19150  
hypothetical protein, 69K - turnip yellow mosaic virus  
C:Species: turnip yellow mosaic virus, TYMV  
C:date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 17-Mar-2000  
C:Accession: S19150  
R:Dreher, T.W.; Bransom, K.L.  
Plant Mol. Biol. 18, 403-406, 1992

A, Title: Genomic RNA sequence of turnip yellow mosaic virus isolate TYMC  
A, Reference number: S19150; MUID:92119261; PMID:1731998  
A, Accession: S19150  
A, Status: preliminary  
A, Molecule type: genomic RNA  
A, Residues: 1-628 <DR>  
C, Cross-references: ENBL:X16378; PIDN:CAA34414.1; PID:g62219  
C, Superfamily: hydroxyproline-rich glycoprotein

```

RESULT 5
G01763
atrophin-1 - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_
C/Accession: G01763
R/Margolis, R.L.
submitted to the EMBL Data Library, March 1995
A/reference number: G08343
A/Accession: G01763
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Restruct: 1-1184 <MAP>
A/Cross-references: EMBL:U23851; NID:G915325; PID:G915326
C/Genetics:
A/Gene: GDB:DRPLA; B37
A/Cross-references: GDB:270336; OMIM:125370
A/Map position: 12p-12p

```



Db 140 SHYLRCSPLDSS--GAGSSRSDDSHSHHTOPSSRTVGNPGNGGFSFSPSGFSEVAPPA 197  
 QY 80 PDR-----LAVPPFAQRERHFRQPTYPYLQHEI-----DLPTTISLSDGEEPPP 125  
 Db 198 PPRNPTASATPP-----PPVPTTQAVYKRRSPALNNRPPAIA-----PPT 241  
 QY 126 YQ--PCTIQ--LRPEQOELNRLNSVRAPPNRTIFDSLMDLSARLGGCPSPSSNGISGISA 181  
 Db 242 ORGNSPVITQGLKPPQQLTQOLKSLNLYPG-----CSGAVVEPPPYLQG--- 290  
 QY 182 TCYSGGRMEGPP-----PTYSVIGH--YFGSFFHQQSGGPPS 219  
 Db 291 ---GAGGAAPPPPSYATSMOSRQSPQSQSDYRKSPSSGIYSATSAGSPS 340

RESULT 10  
 A59249  
 class VII unconventional myosin - slime mold (Dictyostelium discoideum)  
 C:Species: Dictyostelium discoideum  
 C>Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 08-Sep-2000  
 C:Accession: A59249  
 R:Titus, M.A.; Kuspa, A.; Loomis, W.F.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994  
 A:Title: Discovery of myosin genes by physical mapping in Dictyostelium.  
 A:Reference number: Z20873; MUID:95023928; PMID:7937787  
 A:Accession: A59249  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-2357 <TIT>  
 A:CROSS-references: GB:L35321; NID:G6226760; PIDN:AAF06035.1; PID:G6226761  
 A:Experimental source: gene myoI; product class VII unconventional myosin  
 R:Titus, M.A.  
 Curr. Biol. 9, 1297-1303, 1999  
 A:Title: A class VII unconventional myosin is required for phagocytosis.  
 A:Reference number: A59249  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: myoI  
 C:Superfamily: myosin motor domain homology  
 F:16-676/Domain: myosin motor domain homology <MMO>

Query Match 7.9%; Score 107; DB 2; Length 2357;  
 Best Local Similarity 28.1%; Pred. No. 3.3;  
 Matches 43; Conservative 12; Mismatches 56; Indels 42; Gaps 7;

QY 69 IPEQVYAPRPTDLAVPPFAQRERHFRQPTYPYLQHEIDLPTTISLSDGEEPPPQVQ 128  
 Db 1680 VPPQ---PVHPVATLSPP-----MSPTIPNITN--TPPPPPSISDSMSPPQVG 1724

QY 129 PCTQLRDEQOELNRLNSVRAPPNRTIFDSLMDLSARLGGCPSPSSNGISATCYGSGG 188  
 Db 1725 ----MLPPP-----PPSVMGSTKPIEIPSLGIPPPPPSSNSVVPNSPIGS 1767

QY 189 RMEG---PPPTYSEVIGHYFGSFFHQQSGGPP 218  
 Db 1768 PMWGTPPPPTTISV-----HSLNSGNSSTPP 1793

RESULT 11:  
 T33606  
 hypothetical protein F54C4.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T33606  
 R:Rohlfing, T.; Antoniou, B.  
 submitted to the EMBL Data Library, October 1998  
 A:Description: The sequence of C. elegans cosmid F54C4.  
 A:Reference number: Z21376  
 A:Accession: T33606  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1145 <ROH>  
 A:CROSS-references: EMBL:AF099916; PIDN:AAC68776.1; GSPDB:GN00021; CESP:F54C4.3

A:Experimental source: strain Bristol N2; clone F54C4  
 C:Genetics:  
 A:Gene: CESP:F54C4.3  
 A:Map position: 3  
 A:Introns: 57/3; 98/1; 167/3; 193/1; 264/3; 307/3; 357/2; 432/1; 515/3; 564/2; 63/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein F54C4.3

Query Match 7.8%; Score 105; DB 2; Length 1145;  
 Best Local Similarity 21.6%; Pred. No. 2.1;  
 Matches 49; Conservative 25; Mismatches 69; Indels 84; Gaps 8;

QY 86 VPPFAQRERHFRQPTYPYLQHEIDLPTTISLSDGEEPPPQVQCTQLRDEQOELN 144  
 Db 929 IPP--HRQLYQKLDPMFVISHLPSP-----SSRIRIPTRKIQLD 969

QY 145 RESVAPPNRTIFDSLMDLSARLGG-----PCPPSSNGISATCYGSGGRME----- 191  
 Db 970 QESVLPVPPPRRTTSSVVRKRDRFAELRDCPPVEMRAIAAHAIAYDGRKKEORPIVY 1029

QY 192 -----GPPPTYSEVIGHYFGSFFHQQ----- 213  
 Db 1030 VPROKAMAMSTPTAPTTPEDYDGPQHMTTILQHQKQAESEESOWSSEYDDELHEQ 1089

QY 214 -----SSGPPSLLEGTR---LHH---THIAPLESAAIWSKEDKQ 247  
 Db 1090 EHPATSPFPFSLLPHTGSPSEIVHHEEVHMLPEEADLMDDDQ 1136

RESULT 12  
 S01955  
 hypothetical protein, 69K - turnip yellow mosaic virus  
 C:Species: turnip yellow mosaic virus, TYMV  
 C>Date: 21-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-2000  
 C:Accession: S01955  
 R:Morch, M.D.; Boyer, J.C.; Haenni, A.L.  
 Nucleic Acids Res. 16, 6157-6173, 1988  
 A:Title: Overlapping open reading frames revealed by complete nucleotide sequencing of  
 A:Reference number: S01955; MUID:88289359; PMID:3399388  
 A:Accession: S01955  
 A:Status: preliminary  
 A:Molecule type: genomic RNA  
 A:Residues: 1-628 <MOR>  
 A:CROSS-references: EMBL:X07441; NID:G62222; PIDN:CAA30321.1; PID:G62223  
 A:Note: the authors translated the codon ACG for residue 459 as U  
 C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 7.6%; Score 103; DB 2; Length 628;  
 Best Local Similarity 25.3%; Pred. No. 1.6;  
 Matches 58; Conservative 19; Mismatches 80; Indels 72; Gaps 11;

QY 40 RHOQRRREDALSEGCLWPSESTVSGNG-----IPE---POVYAPRPTDLAVPPFAQ 91  
 Db 394 RH---RRPYLLPAPPALPSIAYTSSRGKIHSLPKGALPKGAPP--PPRLPSF--AP 447

QY 92 RERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPQVQCTQLRDEQOELN----- 144  
 Db 448 R-----PQLPLRDLGTPGPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 487

QY 145 -----RESVAPPNRTIFDSLMDLSARLGGCPSPSSNGISATCYGSGGRMEGPPPT 196  
 Db 488 DSDPVLVSVRTEVHAPERRTFMDPEALRSALASLSPSPRS-VGIHTA-----PQT 536

QY 197 YSEVIGHYFGSFFHQQSGGPPSLLEGTRLHHTHIAPLESAAIWSKEDK 245  
 Db 537 ---VLPANPPSPTRHLPTSPFWILQ-----SPVGEDAIVDSEDD 573

RESULT 13  
 T29265  
 hypothetical protein C01G8.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T29265



R.Du, Z.; Gattung, S.  
submitted to the EMBL Data Library, November 1996  
A:Description: The sequence of C. elegans cosmid C01G8.  
A:Reference number: Z20597  
A:Accession: T29265  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1357 <DUZ>  
A:Cross-references: EMBL:U80439; PIDN:AAB37645.1; GSPDB:GN00019; CESP:C01G8.7  
A:Experimental source: strain Bristol N2; clone C01G8  
C:Genetics:  
A:Gene: CESP:C01G8.7  
A:Map position: 1  
A:Introns: 99/1; 488/1; 701/3; 1056/2; 1159/3; 1197/1; 1312/3

Query Match 7.6%; Score 102.5; DB 2; Length 1357;  
Best Local Similarity 21.1%; Pred. No. 4.1;  
Matches 60; Conservative 18; Mismatches 83; Indels 123; Gaps 12;  
QY 59 PSESTVSGNGIPEPQV-----YAPP-----RPTDLAVPPFA----- 90  
DB 343 PQQPPQGGPPGPGQHQHFGYGYGPGAMRPPAGFAPPGAPYGYPPGAPPPAGFH 402  
QY 91 -----QRRERFRQ-----PTYPYLQHEIDLPTTISLSDG 120  
DB 403 PSHQPHQHAQVLAQQRYHQQQHQHQOQQGAGCGRRPPYPYPGGPV--PPGPPQNRM 460  
QY 121 BEPPYPQPCITQLRDPEQQLNRESVR-----APPNRTIFDSDLMSARLGGPCPPSSN 176  
DB 461 PPPPAQG-----APSPGAGSNGKQPRYGTPTAPPSR-----ASATPQPLSST 505  
QY 177 SGISATCGSGRMGPPPTSEVI-----GHYGGSSQF-----QQSGGPPSL 220  
DB 506 MPVAVPSTST-----QPTPTTGSYLANTLATGPAHAPSSMSHAHVISQQOHOYPPGC 560  
QY 221 LEGTLHHTHI-----APLESAATWS 241  
DB 561 IEATATSONQVRKRVYARELINATPRRLIMSLRSLDAAEIAW 604

RESULT 14  
S51796  
vasodilator-stimulated phosphoprotein VASP - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 15-Jul-1995 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S51796  
R:Haffner, C.; Jarchau, T.; Reinhard, M.; Hoppe, J.; Lohmann, S.M.; Walter, U.  
EMBO J. 14, 19-27, 1995  
A:Title: Molecular cloning, structural analysis and functional expression of the proline  
A:Reference number: S51796; MUID:95129547; PMID:7828592  
A:Accession: S51796  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-384 <HAF>  
A:Cross-references: EMBL:Z46388; NID:G624249; PIDN:CAA86522.1; PID:G624250  
C:Keywords: phosphoprotein

Query Match 7.5%; Score 101; DB 2; Length 384;  
Best Local Similarity 26.3%; Pred. No. 1.3;  
Matches 55; Conservative 16; Mismatches 64; Indels 74; Gaps 12;  
QY 47 REDALSSEGCLWPSESTVSGNGIPEPQVYAPRPTDRLAVPPFAQR-ERFHRFQPTYP-Y 104  
DB 96 KEDATQFAAAMASALEAGGGPPPPPPAAP-PTWSVQNGPASEVEQKQKQPGPEH 154  
QY 105 LOHEID-----LPTTISLSDGEEPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFD 158  
DB 155 LERRVSNAGGPPAPP-----ACGPPPPGPPP-----PP----- 183  
QY 159 SDLMDSARLGGPCPPS-SNSGISATCGSGRMGEGPP-----PTYSEVIGHYPGS----- 207  
DB 184 -----GPPPPPGVSLUGSAGAGGAGG---GPPAPPPLTAQGTGGGTGAPGLAA 231

QY 208 -----SFOHQQSSGPP-SLLEGTR 225  
DB 232 AIAGAKLRKVKQBEASGPPVPVPAESTR 260  
RESULT 15  
T28734  
hypothetical protein F26G5.9 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T28734  
R:Sammons, L.; Wohldmann, P.; Beck, C.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid F26G5.  
A:Reference number: Z20516  
A:Accession: T28734  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-944 <SAM>  
A:Cross-references: EMBL:AF022974; PIDN:AAC48041.1; GSPDB:GN00023; CESP:F26G5.9  
A:Experimental source: strain Bristol N2; clone F26G5  
C:Genetics:  
A:Gene: CESP:F26G5.9  
A:Map position: 5  
A:Introns: 12/3; 48/3; 258/3; 327/3; 643/1; 916/2  
Query Match 7.5%; Score 101; DB 2; Length 944;  
Best Local Similarity 27.9%; Pred. No. 3.5;  
Matches 58; Conservative 16; Mismatches 66; Indels 68; Gaps 15;  
QY 61 ESTVSGNGIPEP-----QVYAP-----PRPTD-----RLAVPPPAQRERFHRFQPTYP 104  
DB 339 EPTGSVNQIPASEGGAERVNEVAAPAGRPQPQARNLRGALNHFAQVARNRNQARGVQ 398  
QY 105 LOHEID-LPPTISLSDGEEPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSDLMD 163  
DB 399 QQDLPFLPP-----PPPHQ-----LVMPPP-----AMPPPHQ-----MQ 428  
QY 164 SARL-GGFCPPSSNSGISATCGSGRMGEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLE 222  
DB 429 AQLWGGFPQPMNQPMMAQV---GGR-----PGQY---IMGHPP---PMHHQMQQGPFRMY 477  
QY 223 GTRLHHTHIAPLESAAIWSKEKDKQKH 250  
DB 478 MHQQHH-----EMHQV-----EDQMQRQH 496  
Search completed: December 4, 2003, 19:52:06  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:49:24 ; Search time 21 Seconds  
(without alignments)  
507.730 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MAELFEVQIIIIIVVMVMV.....PLESAAIWSKDKQKHPL 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents At:  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pdp:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pdp:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pdp:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pdp:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pdp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	252	4	US-09-769-482-3
2	1332	98.8	249	4	US-09-769-482-11
3	845.5	62.7	244	4	US-09-769-482-12
4	842.5	62.5	366	4	US-09-769-952A-3
5	791.5	58.7	288	4	US-09-091-952A-4
6	111.5	8.3	1184	4	US-09-266-225D-18
7	111.5	8.3	1185	3	US-09-041-886-23
8	110.5	8.2	732	1	US-08-317-522A-5
9	110.5	8.2	778	1	US-08-438-818A-5
10	110.5	8.2	778	2	US-08-751-965-5
11	110.5	8.2	778	2	US-08-738-975-5
12	110.5	8.2	778	2	US-08-728-626-5
13	110.5	8.2	778	3	US-08-808-599A-5
14	109.5	8.1	1088	4	US-09-233-857-13
15	109.5	8.1	1099	4	US-09-442-100-2
16	109.5	8.1	1099	4	US-08-939-106-2
17	106	7.9	478	4	US-09-252-991A-31455
18	99	7.3	288	4	US-08-396-479B-6
19	97.5	7.2	902	1	US-08-818-823-6
20	97.5	7.2	476	3	US-09-189-035-1
21	96.5	7.2	476	3	US-09-382-086-1
22	96.5	7.2	695	2	US-08-701-240-4
23	96.5	7.2	695	3	US-09-138-236-4
24	96.5	7.2	695	3	US-08-395-580-2
25	95.5	7.1	859	1	PCT-US95-02792-2
26	95.5	7.1	1859	2	US-08-557-139-2
27	95.5	7.1	1243	2	US-08-557-139-2

28	94.5	7.0	913	1	US-08-445-640-4	Sequence 4, Appli
29	94.5	7.0	913	3	US-08-170-558-4	Sequence 4, Appli
30	94.5	7.0	913	3	US-08-447-314-4	Sequence 4, Appli
31	94.5	7.0	913	3	US-08-445-461-4	Sequence 4, Appli
32	94.5	7.0	919	1	US-08-336-343A-2	Sequence 2, Appli
33	94.5	7.0	1242	4	US-09-508-691-1	Sequence 1, Appli
34	93.5	6.9	534	4	US-09-087-134-11	Sequence 11, Appli
35	93.5	6.9	671	3	US-09-121-321-16	Sequence 16, Appli
36	93.5	6.9	671	3	US-08-933-803A-16	Sequence 16, Appli
37	93.5	6.9	1290	3	US-09-150-460B-6	Sequence 6, Appli
38	93	6.9	776	4	US-09-252-991A-28446	Sequence 28446, A
39	93	6.9	1495	4	US-08-522-726B-1	Sequence 1, Appli
40	93	6.9	1495	4	US-09-337-384-1	Sequence 1, Appli
41	92.5	6.9	522	4	US-09-252-991A-28065	Sequence 28065, A
42	92.5	6.9	556	4	US-09-252-991A-17793	Sequence 17793, A
43	92	6.8	518	3	US-09-113-309-19	Sequence 19, Appli
44	92	6.8	518	3	US-09-521-109-19	Sequence 19, Appli
45	92	6.8	518	4	US-09-562-332-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1  
US-09-769-482-3  
; Sequence 3, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995 0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; PRIORITY FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-769-482-3

Query Match	100.0%	Score 1348;	DB 4;	Length 252;
Best Local Similarity	100.0%	Pred. No. 1.4e-118;		
Matches 252;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAELFEVQIIIIIVVMVMVVIITCLLSHYKLSARSFISRHSGRRRDLALSSEGCCLWPS	60	
Db	1	MAELFEVQIIIIIVVMVMVVIITCLLSHYKLSARSFISRHSGRRRDLALSSEGCCLWPS	60	
Qy	61	ESTVSGNGIPEPVYAPPRPTDRLAVPPFAQORERFHRFQPTYVYLQHEIDLPPTISLSDG	120	
Db	61	ESTVSGNGIPEPVYAPPRPTDRLAVPPFAQORERFHRFQPTYVYLQHEIDLPPTISLSDG	120	
Qy	121	EEPPYPQPCPTQLQRDPEQQLLELNRESVRAPPNRTIFDSLDMSARLGGCPPPSSNSGIS	180	
Db	121	EEPPYPQPCPTQLQRDPEQQLLELNRESVRAPPNRTIFDSLDMSARLGGCPPPSSNSGIS	180	
Qy	181	ATCYGSGRMGGPPPTYSVEVGHYPGSSFOHQSSGPPSLLEGLRLHHHTHIAPLESAAIW	240	
Db	181	ATCYGSGRMGGPPPTYSVEVGHYPGSSFOHQSSGPPSLLEGLRLHHHTHIAPLESAAIW	240	
Qy	241	SKEYDKQKHPL	252	
Db	241	SKEYDKQKHPL	252	

```
RESULT 2
US-09-769-482-11
; Sequence 11, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: POYNUCLEOTIDE ARRAY
; CURRENT FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2000-01-28
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-11

Query Match      98.8%; Score 1332; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 4.2e-117;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 AAEFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSEGCWPSE 61
DB  1 AAEFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSEGCWPSE 60

QY  62 STVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFQTPYQLQHEIDLPTTISLSDGE 121
DB  61 STVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFQTPYQLQHEIDLPTTISLSDGE 120

QY  122 EPPYQGPCTQLRDPQOOLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 181
DB  121 EPPYQGPCTQLRDPQOOLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180

QY  182 TCYSGGGRMEGPPPTYSVIGHYCGSSFHQOQSSGPPSLLEGLRLHHTHAPLESAAIWS 241
DB  181 TCYSGGGRMEGPPPTYSVIGHYCGSSFHQOQSSGPPSLLEGLRLHHTHAPLESAAIWS 240

QY  242 KEKDKQKG 250
DB  241 KEKDKQKG 249

RESULT 3
US-09-769-482-12
; Sequence 12, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
```

```
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-12

Query Match      62.7%; Score 845.5; DB 4; Length 244;
Best Local Similarity 67.6%; Pred. No. 1.6e-71;
Matches 169; Conservative 21; Mismatches 53; Indels 7; Gaps 4;

QY  2 AAEFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSEGCWPSE 61
DB  1 AAEFVQIIIVVMVMVVITCLSHYKLSARSFISRHSGRRRDLSEGCWPSE 60

QY  62 STVSGNGIPEQVYAPRPTDRLAVPPAQRERFHRFQTPYQLQHEIDLPTTISLSDGE 121
DB  61 SAAPRLGASE--IMHAPRSRDRFTAPFIQDRFSRFOPTYPYQVQHEIDLPTTISLSDGE 118

QY  122 EPPYQGPCTQLRDPQOOLNRESVRAPNRTIFDSLDMSARL--GGPCPPSSNSGIS 180
DB  119 EPPYQGPCTQLRDPQOOLNRESVRAPNRTIFDSLDIDIAMYSGPCPPSSNSGIS 178

QY  181 ATCYSGGGRMEGPPPTYSVIGHYCGSSFHQOQSSGPPSLLEGLRLHHTHAPLESAAIWS 240
DB  179 ASTCSSNGRMEGPPPTYSVEMGHHPGASFLHHQRS--NAHRGSRLOFQQ--NNAESTTVP 234

QY  241 SKEKDKQKG 250
DB  235 IKGDKRKP 244

RESULT 4
US-09-091-952A-3
; Sequence 3, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
```

TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-OCT-1996  
PCT/US97/19381  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-297100US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-091-952A-3

Query Match 62.5%; Score 842.5; DB 4; Length 306;  
Best Local Similarity 67.2%; Pred. No. 4.1e-71;  
Matches 168; Conservative 22; Mismatches 53; Indels 7; Gaps 4;  
QY 2 ALEFVQIIIVVMVMMVITCLSHYKLSARSFISRHSGRRREDALSEGCLWPSE 61  
DB 61 SELEPAQIIIVVMVMMVITCLSHYKLSARSFISRHSGRRREDALSEGCLWPSE 120  
QY 62 STVSGNGIPEQVYAPPRDRLAVPPFAQRERFHRFQTPYVQHEIDLPTTSLSDGE 121  
DB 121 SAAPRLGASE--IMHAPSRDRFTAPSFQDRFRFQTPYVQHEIDLPTTSLSDGE 178  
QY 122 EPPYQGPCTQLDRDPEQLELNRESVRAPNRTIFDSLDMSARL--GGPCPPSSNSGIS 180  
DB 179 EPPYQGPCTQLDRDPEQLELNRESVRAPNRTIFDSLDIDIAMYSGGPCPPSSNSGIS 238  
QY 181 ATCYSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSGPPSLLEGTRLHHHTHAPLESAAIW 240  
DB 239 ASTCSSNGRMEGPPPTTYSEVGMHGHFGASFLHHQRS---NAHRGSLQFQQ--NNAESTIVP 294  
QY 241 SKEDKQKQH 250  
DB 295 IKGDKRKPEN 304

RESULT 5  
US-09-091-952A-4  
Sequence 4, Application US/09091952A  
Patent No. 6458532  
GENERAL INFORMATION:  
APPLICANT: Detera-Wadleigh, Sevilla D.  
Gershon, Elliot S.  
Badner, Judith A.  
Goldin, Lynn R.  
Berrettini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Esterling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999

CLASSIFICATION: <Unknown>  
PRIORITY DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-OCT-1996  
PCT/US97/19381  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-297100US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 288 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: mat peptide  
LOCATION: 1--288  
OTHER INFORMATION: Clone 22 isoform 2  
alternatively spliced  
protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-091-952A-4  
Query Match 58.7%; Score 791.5; DB 4; Length 288;  
Best Local Similarity 64.0%; Pred. No. 2.3e-66;  
Matches 160; Conservative 21; Mismatches 44; Indels 25; Gaps 4;  
QY 2 ALEFVQIIIVVMVMMVITCLSHYKLSARSFISRHSGRRREDALSEGCLWPSE 61  
DB 61 SELEPAQIIIVVMVMMVITCLSHYKLSARSFISRHSGRRREDALSEGCLWPSE 110  
QY 62 STVSGNGIPEQVYAPPRDRLAVPPFAQRERFHRFQTPYVQHEIDLPTTSLSDGE 121  
DB 111 -----PQIMHAPSRDRFTAPSFQDRFRFQTPYVQHEIDLPTTSLSDGE 160  
QY 122 EPPYQGPCTQLDRDPEQLELNRESVRAPNRTIFDSLDMSARL--GGPCPPSSNSGIS 180  
DB 161 EPPYQGPCTQLDRDPEQLELNRESVRAPNRTIFDSLDIDIAMYSGGPCPPSSNSGIS 220  
QY 181 ATCYSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSGPPSLLEGTRLHHHTHAPLESAAIW 240  
DB 221 ASTCSSNGRMEGPPPTTYSEVGMHGHFGASFLHHQRS---NAHRGSLQFQQ--NNAESTIVP 276  
QY 241 SKEDKQKQH 250  
DB 277 IKGDKRKPEN 286

RESULT 6  
US-09-266-225D-18  
Sequence 18, Application US/09266225D  
Patent No. 6573364  
GENERAL INFORMATION:  
APPLICANT: Nandabalan, Krishan  
APPLICANT: Kingsmore, Stephen  
APPLICANT: Tchernev, Velizar  
TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak Syndrome (HPS) Protein Complexes and HPS Protein-TITLE OF INVENTION: Interacting Proteins  
FILE REFERENCE: 15966-523  
CURRENT APPLICATION NUMBER: US/09/266,225D  
CURRENT FILING DATE: 1999-03-10  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 18  
LENGTH: 1184



RESULT 9  
US-08-439-818A-5  
; Sequence 5, Application US/08439818A  
; Patent No. 5654145  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/439,818A  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/317,522  
; FILING DATE: 04-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 1563  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 778 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-439-818A-5

Query Match 8.2%; Score 110.5; DB 1; Length 778;  
Best Local Similarity 24.5%; Pred. No. 0.054;  
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;  
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQERFHRFQPTY--PYLQHEIDLPTTI 115  
DB 498 SGLPKPCLPECGEPQPCPAEPG-----PP-----EAFCRSEPEIPEPSLQLEVPPEPY 548  
QY 116 SLSDGEPPPYQGCTQLQDRPE-----QOLELNRESVRAPP-----NRTIFDSDL 162  
DB 549 PPA---EPPLESCCRSEPEIPESSRQEQLEVPPEPCPAEPPLSEYCRIEPEISRRQ 605  
QY 163 DSARLGCPSPSSNGISATCYGSGRMGEGPPPTYSEVIGHYPGSSFOHQ--QSSGPPSL 220  
DB 606 EQLEVPPEPCPAEPGLQPSQTQSG-PPGCPRVE--LGASEPCTLEHRSLESLLPCC 662  
QY 221 LEGTLHHTHTAPLESAAIWSKEDKQKHPL 252  
DB 663 -----SQWAPATSLIFSSQ-----HPL 680

RESULT 10  
US-08-751-965-5  
; Sequence 5, Application US/08751965  
; Patent No. 5858360  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting

; TITLE OF INVENTION: Proteins  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/751,965  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/439,818  
; FILING DATE: 12-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 2252  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 778 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-751-965-5  
Query Match 8.2%; Score 110.5; DB 2; Length 778;  
Best Local Similarity 24.5%; Pred. No. 0.054;  
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;  
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQERFHRFQPTY--PYLQHEIDLPTTI 115  
DB 498 SGLPKPCLPECGEPQPCPAEPG-----PP-----EAFCRSEPEIPEPSLQLEVPPEPY 548  
QY 116 SLSDGEPPPYQGCTQLQDRPE-----QOLELNRESVRAPP-----NRTIFDSDL 162  
DB 549 PPA---EPPLESCCRSEPEIPESSRQEQLEVPPEPCPAEPPLSEYCRIEPEISRRQ 605  
QY 163 DSARLGCPSPSSNGISATCYGSGRMGEGPPPTYSEVIGHYPGSSFOHQ--QSSGPPSL 220  
DB 606 EQLEVPPEPCPAEPGLQPSQTQSG-PPGCPRVE--LGASEPCTLEHRSLESLLPCC 662  
QY 221 LEGTLHHTHTAPLESAAIWSKEDKQKHPL 252  
DB 663 -----SQWAPATSLIFSSQ-----HPL 680  
RESULT 11  
US-08-738-975-5  
; Sequence 5, Application US/08738975  
; Patent No. 5880267  
; GENERAL INFORMATION:  
; APPLICANT: Fukuda, Michiko N.  
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,975  
FILING DATE: herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/439,818  
FILING DATE: 05-Dec-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 2251  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 778 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-738-975-5

Query Match 8.2%; Score 110.5; DB 2; Length 778;  
Best Local Similarity 24.5%; Pred. No. 0.054;  
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;

QY 67 NGIP-----EPQVYAPRPTDRLAVPPAQRERHFRQPTY--PYLQHEIDLPTI 115  
Db 498 SGLPKPCLPCEGEPQCPPEAG-----PP-----EAFCRSEPIPEPSLQQLVPEPY 548  
QY 116 SLSDGEEPPYQGCCTQLRDPE---QQLNLRNRESVRAPP-----NRTIFDSLM 162  
Db 549 PPA---EPRPLSCCRSEPIESSRQQLVPEPCPPAEPPLSYCRIEPIESSRQ 605  
QY 163 DSARLGGCPSPSSNGISATCYGSGRMEGPPPTVSEVIGHYPGSSFOHQ--QSSGPPSL 220  
Db 606 EQLEVPCEPCPAEPGLOPSTQSG--PPGCPRVE--LGASEPCTLEHRSLSLPPCC 662  
QY 221 LEGTLRHHTHIAPLSAATWSKEDKQKHPL 252  
Db 663 -----SQWAPATSLIFSSQ-----HPL 680

RESULT 12  
US-08-728-626-5  
Sequence 5, Application US/08/28626  
Patent No. 5910451  
GENERAL INFORMATION:  
APPLICANT: Fukuda, Michiko N.  
TITLE OF INVENTION: Trophinin and Trophinin-Assisting  
PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,626  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/439,818  
FILING DATE: 12-MAY-1995  
APPLICATION NUMBER: US 08/317,522  
FILING DATE: 04-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 1563  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 778 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-728-626-5

Query Match 8.2%; Score 110.5; DB 2; Length 778;  
Best Local Similarity 24.5%; Pred. No. 0.054;  
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;

QY 67 NGIP-----EPQVYAPRPTDRLAVPPAQRERHFRQPTY--PYLQHEIDLPTI 115  
Db 498 SGLPKPCLPCEGEPQCPPEAG-----PP-----EAFCRSEPIPEPSLQQLVPEPY 548  
QY 116 SLSDGEEPPYQGCCTQLRDPE---QQLNLRNRESVRAPP-----NRTIFDSLM 162  
Db 549 PPA---EPRPLSCCRSEPIESSRQQLVPEPCPPAEPPLSYCRIEPIESSRQ 605  
QY 163 DSARLGGCPSPSSNGISATCYGSGRMEGPPPTVSEVIGHYPGSSFOHQ--QSSGPPSL 220  
Db 606 EQLEVPCEPCPAEPGLOPSTQSG--PPGCPRVE--LGASEPCTLEHRSLSLPPCC 662  
QY 221 LEGTLRHHTHIAPLSAATWSKEDKQKHPL 252  
Db 663 -----SQWAPATSLIFSSQ-----HPL 680

RESULT 13  
US-08-808-599A-5  
Sequence 5, Application US/08808599A  
Patent No. 611089  
GENERAL INFORMATION:  
APPLICANT: Fukuda, Michiko N.  
TITLE OF INVENTION: Trophinin, Trophinin-Assisting  
PROTEINS AND METHODS TO INHIBIT IMPLANTATION  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,599A  
FILING DATE: 28-FEB-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/317,522  
FILING DATE: 04-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/439,818  
FILING DATE: 12-MAY-1995  
ATTORNEY/AGENT INFORMATION:





Search completed: December 4, 2003, 19:52:39  
Job time : 22 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:44:54 ; Search time 18 Seconds  
(without alignments)  
658.374 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MAELEFVQIIIVVMVMV.....PLESAATWSKEDKQKHPL 252

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1340	99.4	287	1 TMPEP HUMAN	Q969W9 homo sapien
2	1172.5	87.0	260	1 TMPEP MOUSE	Q947R2 mus musculus
3	842.5	62.5	306	1 G8R1 HUMAN	Q15165 homo sapien
4	114	8.5	1509	1 G8R1 HUMAN	Q947R2 mus musculus
5	113.5	8.4	395	1 Q91KAS ARABIDOPSIS	Q91KAS arabidopsis
6	113	8.4	628	1 V70K TYMVC	P28478 turnip yell
7	111.5	8.3	1185	1 DRPL HUMAN	P54259 homo sapien
8	110.5	8.2	778	1 TAST HUMAN	Q12815 homo sapien
9	104	7.7	830	1 SREC HUMAN	Q14162 homo sapien
10	103	7.6	628	1 V70K TYMV	P10357 turnip yell
11	101	7.5	384	1 VASP CANFA	P50551 canis fami
12	100.5	7.5	1183	1 DRPL RAT	P54258 rattus norv
13	100	7.4	628	1 V70K TYMVA	P20131 turnip yell
14	100	7.4	2440	1 NCR1 HUMAN	O75376 homo sapien
15	99.5	7.4	929	1 EBN6 EBV	P03204 epstein-bar
16	98	7.3	1259	1 AUT2 HUMAN	Q8WXX7 homo sapien
17	98	7.2	2161	1 SHK1 HUMAN	Q9Y566 homo sapien
18	97.5	7.2	902	1 NRC4 HUMAN	Q14934 homo sapien
19	96	7.1	620	1 EXTN TOBAC	P13983 nicotiana t
20	96	7.1	1729	1 TRAP HUMAN	Q9C0C2 homo sapien
21	96	7.1	1844	1 POLR TYMVA	P01028 turnip yell
22	95.5	7.1	707	1 SFQ HUMAN	P23246 homo sapien
23	95.5	7.1	859	1 M3KC HUMAN	Q12852 homo sapien
24	95	7.0	896	1 CYRB MOUSE	P26955 mus musculus
25	94.5	7.0	276	1 PRPL HUMAN	P10162 homo sapien
26	94.5	7.0	466	1 Y137 CAEEL	P34428 caenorhabdi
27	94.5	7.0	913	1 DR1 HUMAN	Q08345 h epithelia
28	94.5	7.0	1242	1 IRS1 HUMAN	P35568 homo sapien
29	94.5	7.0	1790	1 SEPA EMENI	P78621 emerice
30	93.5	6.9	671	1 Z282 HUMAN	Q9UDV7 homo sapien
31	93.5	6.9	910	1 DR1 RAT	Q63474 rattus norv
32	93.5	6.9	2083	1 DYSF MOUSE	Q9ESD7 mus musculus
33	93	6.9	244	1 YPHF_BACSU	P39911 bacillus su

#### RESULT 1

ID	TMPEP_HUMAN	STANDARD;	PRT;	287 AA.
AC	Q969W9; Q96B72; Q9UCD3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Transmembrane prostate androgen-induced protein (Solid tumor-associated 1 protein).			
GN	TMPEP1 OR PMP1 OR STAG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE=20334621; PubMed=10873380;			
RA	Xu L.L., Shanmugam N., Segawa T., Seestherenn I.A., McLeod D.G.,			
RA	Moul J.W., Srivastava S.;			
RT	"A novel androgen-regulated gene, PMP1, located on chromosome 20q13 exhibits high level expression in prostate."			
RL	Genomics 66:257-263 (2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=21453682; PubMed=11568975;			
RA	Rae F.K., Hooper J.D., Nicol D.L., Clements J.A.;			
RT	"Characterization of a novel gene, STAG1/PMP1, upregulated in renal cell carcinoma and other solid tumors."			
RL	Mol. Carcinog. 32:44-53 (2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carter C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Ellington A.G., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Graham D.V., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Hammond S., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Huckle E., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaeslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,			
RA	Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			

Not prior.

RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20.";  
RL Nature 414:865-871(2001).  
RN [4]  
RP SEQUENCE OF 8-287 FROM N.A. (ISOFORM 1).  
RX TISSUE=Kidney;  
RA MEDLINE=2238857; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q969W9-1; Sequences=Displayed;  
CC Name=2;  
CC IsoId=Q969W9-2; Sequences=VSP 006438;  
CC Notes=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Highest expression in prostate. Also expressed  
CC in ovary.  
CC -!- INDUCTION: By androgen.  
CC -!- SIMILARITY: BELONGS TO THE TMEM16 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF224278; RA06322.1; -;  
DR EMBL: AF305616; AAL18781.1; -;  
DR EMBL: AF305426; AAL09357.1; -;  
DR EMBL: AL035541; CAB55862.1; -;  
DR EMBL: BC015918; AAL15918.1; ALT\_INIT.  
DR Genew: HGNC:14107; TMEM16.  
DR MIM: 606564; -;  
DR GO: GO:0016021; C: integral to membrane; NAS.  
DR GO: GO:0030521; P: androgen receptor signaling pathway; NAS.  
DR KW Transmembrane; Alternative splicing.  
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 41 63 POTENTIAL.  
FT DOMAIN 64 287 CYTOPLASMIC (POTENTIAL).  
FT SITE 158 161 WW-BINDING (POTENTIAL).  
FT SITE 229 232 WW-BINDING (POTENTIAL).  
FT VARSPIC 1 37 MRLMGVNSTAAAGQPNVSCNCKRSLFQSMBIT ->  
FT MA (in isoform 2).  
FT /FTID=VSP\_006438.  
SQ SEQUENCE 287 AA; 31609 MW; 6103473561AE08DA CRC64;  
Query Match 99.4%; Score 1340; DB 1; Length 287;  
Best Local Similarity 99.2%; Pred. No. 9.7e-94;  
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAELFEVQIIIIIVMMVMVVVITCLSHYKLSARSFISRHSQGRRRDALSSEGLWPS 60  
DB : |||||  
36 ITELEFVQIIIIIVMMVMVVVITCLSHYKLSARSFISRHSQGRRRDALSSEGLWPS 95  
QY 61 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQRERPHRFQPTYPYQLQHEIDLPTTSLSDG 120  
DB : |||||  
96 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQRERPHRFQPTYPYQLQHEIDLPTTSLSDG 155  
QY 121 EEPYPYQPCITQLRDPQQLQELNRESVRAPPNRTIFSDLMDSARLGGPCPPSSNGIS 180  
DB : |||||  
156 EEPYPYQPCITQLRDPQQLQELNRESVRAPPNRTIFSDLMDSARLGGPCPPSSNGIS 215  
QY 181 ATCVSGSGRMGGPPTVSEVGHYPGSGFQHQSGGPPSLLEGTRLHHTHAPLESAAIW 240  
DB : |||||  
216 ATCVSGSGRMGGPPTVSEVGHYPGSGFQHQSGGPPSLLEGTRLHHTHAPLESAAIW 275  
QY 241 SKERDKQKQHPL 252  
DB : |||||  
276 SKERDKQKQHPL 287  
RESULT 2  
TMEM16 MOUSE  
ID TMEM16 MOUSE STANDARD; PRT; 260 AA.  
AC Q9D7R2; Q9EQH9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Transmembrane prostate androgen-induced protein (Nedd4 WW domain-  
DE binding protein 4).  
GN TMEM16 OR N4WBPA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Stomach;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada S.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection.";  
RT Nature 409:685-690(2001).  
RL [2]  
RP SEQUENCE OF 59-260 FROM N.A.  
RX MEDLINE=20498735; PubMed=11042109;  
RA Jolliffe C.N., Harvey K.F., Haines B.P., Parasivam G., Kumar S.;  
RT "Identification of multiple proteins expressed in murine embryos as  
RT binding partners for the WW domains of the ubiquitin-protein ligase  
RT Nedd4.";  
RL Biochem. J. 351:557-565(2000).  
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4.  
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE TMEM16 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glioma tumor suppressor candidate region gene 1 protein.  
 GN GLTSCR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20175430; PubMed=10708517;  
 RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,  
 RA Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,  
 RA Scheithauer B.W., Louis D.N., Jenkins R.B.;  
 RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor  
 RT region."; Genomics 64:44-50(2000).  
 RL SEQUENCE OF 1509 AA; 152991 MW; 7C5144F443CE6821 CRC64;  
 CC -!- TISSUE SPECIFICITY: Expressed at moderate levels in heart, brain,  
 CC placenta, skeletal muscle, and pancreas, and at lower levels in  
 CC lung, liver, and kidney.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF182077; AAF62874.1; -;  
 CC Genew; HGNC:4332; GLTSCR1.  
 CC MIM; 605690; -;  
 CC DOMAIN 37 45 POLY-GLY.  
 CC DOMAIN 884 889 POLY-PRO.  
 CC DOMAIN 1214 1225 POLY-SER.  
 CC DOMAIN 1282 1286 POLY-PRO.  
 CC DOMAIN 1294 1304 POLY-PRO.  
 CC SEQUENCE 1509 AA; 152991 MW; 7C5144F443CE6821 CRC64;  
 SQ  
 Query Match 8.5%; Score 114; DB 1; Length 1509;  
 Best Local Similarity 26.9%; Pred. No. 0.53;  
 Matches 49; Conservative 20; Mismatches 63; Indels 50; Gaps 10;  
 QY 69 IPEQVYAPP-RPTDR---LAVPFAQRERFH-----RFQTPY---YLOHEIDLPTTIS 116  
 Db 740 LPSHPHTPPSPRPSQSVSRPP--SEPLHPCPPQAPPLGIFVIOQLGVPPPAS 797  
 QY 117 LSDGEEPPYQGFCTQLRDPEQQLNRESVRAAPPNRTIFDSLDMSARLGGPCPPSSN 176  
 Db 798 NPAPTAPGPPQLRPQSPREGPL-----PPAPHL-----PPSST 833  
 QY 177 SGISATCYSGGORMEGPPPTTSEVIGHYPGSFQHQSSGPPSLLEGTRLHHTHTAPLES 236  
 Db 834 SSAVASSETSSRL--PAPTPSDFQLQFPSPGQPHKSPPTPL-----HLVP-EP 881  
 QY 237 AA 238  
 Db 882 AA 883  
 RESULT 5  
 UMPI\_ARATH STANDARD; PRT; 395 AA.  
 AC Q9LKA5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Unknown mitochondrial protein At3g15000.  
 GN AT3G15000 OR K15W2.14.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.  
 OX NCBI\_TaxID=3702;

RN SEQUENCE FROM N.A.  
 RP STRAIN=cv. Columbia;  
 RX MEDLINE=20363099; PubMed=10907853;  
 RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
 RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC  
 RT and BAC clones."; DNA Res. 7:217-221(2000).  
 RL DNA Res. 7:217-221(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
 RT SSP consortium (Isak/Stanford/PCEC)."; Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 169-179, AND SUBCELLULAR LOCATION.  
 RC TISSUE=Leaf, and Stem;  
 RX MEDLINE=21608403; PubMed=11743114;  
 RA Krufft V., Eubel H., Jaensch L., Werhahn W., Braun H.-P.;  
 RT "Proteomic approach to identify novel mitochondrial proteins in  
 RT Arabidopsis."; Plant Physiol. 127:1694-1710(2001).  
 RL Arabidopsis.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AP000370; BAA97063.1; -;  
 CC EMBL; AF428427; AAL16196.1; -;  
 CC Mitochondrion.  
 KW Mitochondrion.  
 SQ SEQUENCE 395 AA; 42869 MW; D15E5CCC89E2A33C CRC64;  
 Query Match 8.4%; Score 113.5; DB 1; Length 395;  
 Best Local Similarity 23.7%; Pred. No. 0.13;  
 Matches 46; Conservative 12; Mismatches 61; Indels 75; Gaps 8;  
 QY 45 RRREDALSSGCLWPESTVSGNGIPEQV--YAPPRPTDLAVPPFAQRERHFRQPTY 102  
 Db 232 RRRENMAAGP-----PPQPPMGPPPPPHIGSGAPPPPHMGSGAPP-----P 274  
 QY 103 PYLQHEIDLPTTISLSDGEEPPYQGPCTQLRDPEQQLNRESVRAAPPNRTIFDSLDL 162  
 Db 275 PHMGQNYGPPPPNNMGGRHPPPYGAP-----PQN----- 304  
 QY 163 DSARLGGPCPPSSNGISATCYGS-----GGRMEGPPPTTYSEVI-----G 202  
 Db 305 ---NMGGPRPQNYGGTTPPNYGGAPPANNMGAPPNNYGG---GPPPYGAVPPPPQYGG 358  
 QY 203 HYPGSSFFHQHQSSG 216  
 Db 359 APPQNNNYQQQSG 372  
 RESULT 6  
 V70K\_TYMWV  
 ID V70K\_TYMWV STANDARD; PRT; 628 AA.  
 AC P28478;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-DEC-1992 (Rel. 24, Last annotation update)  
 DE 69 kDa protein.  
 OS Turnip yellow mosaic virus (isolate TYMC).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.  
 OX NCBI\_TaxID=31751;  
 RN [1]  
 RP SEQUENCE FROM N.A.







"Expression cloning of a novel scavenger receptor from human endothelial cells.";  
 J. Biol. Chem. 272:31217-31220(1997).  
 [2]

SEQUENCE FROM N.A.

RA MEDLINE=22086180; PubMed=11978792;

ADACHI H., Tsujimoto M.;

"Characterization of the human gene encoding the scavenger receptor expressed by endothelial cell and its regulation by a novel transcription factor, endothelial zinc finger protein-2.";  
 J. Biol. Chem. 277:24014-24021(2002).  
 [3]

SEQUENCE FROM N.A.

RP TISSUE=Bone marrow;

RC MEDLINE=96127530; PubMed=8590280;

RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;

"Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1.";  
 DNA Res. 2:167-174(1995).  
 [3]

FUNCTION: Mediates the binding and degradation of acetylated low

density lipoprotein (Ac-LDL). Mediates heterophilic interactions, suggesting a function as adhesion protein (By similarity).

!- SURUNIT: Heterophilic interaction with SREC2 via its extracellular domain. The heterophilic interaction is suppressed by the presence of ligand such as Ac-LDL (By similarity).

!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).

!- TISSUE SPECIFICITY: Endothelial cells.

!- SIMILARITY: Contains 6 EGF-like domains.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL; D86864; BAA24070.1; -.

DR EMBL; AB052946; BAC02692.1; -.

DR EMBL; D63483; BAA09770.1; -.

DR HSSP; P01180; 2EN2. SCARF1.

DR Genew; HGNC:16820; SCARF1.

DR GO; GO:0016021; C:integral to membrane; IDA.

DR GO; GO:0030169; F:low-density lipoprotein binding activity; IDA.

DR GO; GO:0004888; F:transmembrane receptor activity; TAS.

DR GO; GO:0045192; P:low-density lipoprotein catabolism; TAS.

DR GO; GO:0006898; P:receptor mediated endocytosis; TAS.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR002049; Laminin\_EGF.

DR PRINTS; PR00011; EGF\_LAMININ.

DR SMART; SM00181; EGF\_5.

DR PROSITE; PS00022; EGF\_1; 6.

DR PROSITE; PS01186; EGF\_2; 6.

KW Cell adhesion; Receptor; Repeat; Signal; Transmembrane;

KW EGF-like domain; Glycoprotein.

FT SIGNAL 1 19

FT CHAIN 20 830

FT DOMAIN 20 421

FT TRANSMEM 422 442

FT DOMAIN 443 830

FT DOMAIN 53 87

FT DOMAIN 95 130

FT DOMAIN 155 191

FT DOMAIN 215 249

FT DOMAIN 302 339

FT DOMAIN 351 382

FT DOMAIN 476 620

FT DOMAIN 622 798

FT DOMAIN 431 438

FT DOMAIN 57 69

FT DISULFID 53 75

FT POTENTIAL.

FT

FT

FT

FT

FT

FT

FT DISULFID 77 86

FT DISULFID 99 111

FT DISULFID 105 118

FT DISULFID 120 129

FT DISULFID 159 172

FT DISULFID 165 179

FT DISULFID 181 190

FT DISULFID 219 230

FT DISULFID 225 237

FT DISULFID 239 248

FT DISULFID 306 319

FT DISULFID 313 326

FT DISULFID 329 338

FT DISULFID 355 363

FT DISULFID 358 370

FT DISULFID 372 381

FT CARBOHYD 289 289

FT CARBOHYD 382 382

FT CARBOHYD 393 393

FT CONFLICT 662 662

SQ SEQUENCE 830 AA; 87430 MW; F560D9E1AA64D779 CRC64;

Query Match 7.7%; Score 104; DB 1; Length 830;

Best Local Similarity 22.6%; Pred. No. 1.5;

Matches 60; Conservative 32; Mismatches 95; Indels 78; Gaps 11;

QY 36 SFISRHSGQRREDALSSG-----CLWP-----SESTVSGNGIPEP 72

DB 508 SFIFPPSAGWATDDSFSDPSGEADVPACVFPQGMVFPVQAQGSSEASLAAGAFPPP 567

QY 73 QVYAPPRTDLAVPPFAQRERFHRFQTPYLOHEIDLPTTISLSDGEEPPPY----QG 128

DB 568 EDASTP-----FAIPRTSSSLAKR-----PSVSPAECTKPAQSRSSG 607

QY 129 PCTQLRDPQOQLNRESVPAPNRTIFDSDLNDSARLGGPCPPSSNSGISATCYSGG 188

DB 608 ELSSPLRKPKR---LSRGAQSGPEGREABESTGPDEAEAPESFPAAAGPGDSAT----GH 660

QY 189 RMEGPPPTYSSEVIGHYPGSSSFHQOQSSGP-----PSLLEG-TLHHHTHIAPLESA 237

DB 661 RR--PPLGSRIVAEHVEAGSVQESSGPVTIYMLACKPRGSEGPVRSVRFHFGSKG 718

QY 238 AIMSKEK-----DKOKGHP 251

DB 719 QAEAKVKRAIKPQPPQALNRKKGSP 743

RESULT 10

V70K TYMV

ID V70K TYMV STANDARD; PRT; 628 AA.

AC P10357;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-AUG-1992 (Rel. 23, Last annotation update)

DE 69 kDa protein.

OS Turnip yellow mosaic virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.

OX NCBI\_taxID=12154;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=88289359; PubMed=3399388;

RA Morch M.D., Boyer J.C., Haenni A.L.;

"Overlapping open reading frames revealed by complete nucleotide

sequencing of turnip yellow mosaic virus genomic RNA.";

Nucleic Acids Res. 16:6157-6173(1988).

CC !- FUNCTION: NOT KNOWN.

CC !- SIMILARITY: TO 65 TO 70 kDa PROTEIN FROM OTHER TYMOVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

use by and for commercial



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U31777; AAA80337.1; -  
EMBL; X89453; CA61823.1; -  
InterPro; IPR002951; Atrophin.  
Pfam; PF03154; Atrophin-1; 2.  
PRINTS; PR01222; ATROPHIN.

DOMAIN 165 171 POLY-PRO.  
DOMAIN 303 306 POLY-PRO.  
DOMAIN 377 383 POLY-SER.  
DOMAIN 387 391 POLY-SER.  
DOMAIN 440 446 POLY-PRO.  
DOMAIN 477 480 POLY-HIS.  
DOMAIN 481 489 POLY-GLN.  
DOMAIN 502 505 POLY-PRO.  
DOMAIN 562 572 POLY-SER.  
DOMAIN 702 705 POLY-PRO.  
DOMAIN 455 455 N -> S (IN REF. 2).  
CONFLICT 594 594 F -> L (IN REF. 2).  
CONFLICT 689 689 P -> R (IN REF. 2).  
CONFLICT 717 717 T -> M (IN REF. 2).  
CONFLICT 737 737 A -> V (IN REF. 2).  
CONFLICT 965 965 MISSING (IN REF. 2).  
SEQUENCE 1183 AA; 124778 MW; 7FB9928DCADF9B1F CRC64;

Query Match 7.5%; Score 100.5; DB 1; Length 1183;  
Best Local Similarity 22.6%; Pred. No. 4.1;  
Matches 35; Conservative 14; Mismatches 45; Indels 51; Gaps 7;  
QY 100 PTPYVQLQHEIDLPTTISLSDGEEPPYQGPCTQLQRDPEQQLNRESVRAPPNRTIFDS 159  
DB 404 PSYHP-----SFPPTSVSNQPKYTOP-----SLPSQAVWSQ 438  
QY 160 DLMDSARLGRCPPSSNGISATCYGSGRMGCP-PPYSEVIGHYPPSSFHQQ----- 213  
DB 439 -----GPPPPPPYX-----RLLPNNTHGPPPTGGQSTAHPPAPAHHHHQQQQQP 486  
QY 214 -----SSGP-----PSLEGLRLHHTH 230  
DB 487 QPQPOQHGHGNSGPPPGAYPHPLESSNSHAAH 521

RESULT 13  
ID V70K TYMVA  
AC P20131; STANDARD; PRT; 628 AA.  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE 69 kDa protein.  
OS Turnip yellow mosaic virus (Australian isolate).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.  
OX NCBI\_TaxID=12155;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90021184; PubMed=2800335;  
RA Keese P., Mackenzie A., Gibbs A.;  
RT "Nucleotide sequence of the genome of an Australian isolate of turnip yellow mosaic tymovirus.";  
RL Virology 172:536-546(1989).  
CC -!- FUNCTION: NOT KNOWN.  
CC -!- SIMILARITY: TO 65 TO 70 kDa PROTEIN FROM OTHER TYMOVIRUSES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; J04373; AAA46591.1; -  
DR PIR; JQ0110; JQ0110.  
DR InterPro; IPR004935; Tymo 45 70kDa.  
DR Pfam; PF03251; Tymo 45kd 70kd; 1.  
SQ SEQUENCE 628 AA; 68740 MW; 67CD342A09161D96 CRC64;

Query Match 7.4%; Score 100; DB 1; Length 628;  
Best Local Similarity 25.0%; Pred. No. 2.2;  
Matches 55; Conservative 19; Mismatches 95; Indels 54; Gaps 10;

QY 40 RHOGRREDALSSGCLWPSSTVSGNG-----IPFQVYAPRPTDRLAVPPF 89  
DB 394 RH-----RRSHPLLPNPPAALPIAYTSGRGKIHLSPKGLPKGPPPPRRLPSPATPPQ 450  
QY 90 AQERFHRFQPTYPVQLQHEIDLPTTISLSDGEE-----PPPYQGPCTQLQRDPEQQLNLR 145  
DB 451 SPLRDLGR-TESFP-----TPKTSRATESCIAAPTDIAPL-----DSDPVLVS-R 496  
QY 146 ESVRAPNRTIFDLSMDLSARLGRCPPSSNGISATCYGSGRMGCPPTTYSEVIGHYP 205  
DB 497 TEVHAPERTFMDPEALRAALAGLSPPLS-VGIIRTSQTVLPTNSPSPT----- 546  
QY 206 GSSFQHQSSGPPSLLEGLRLHHTHIAPLSAIAIWSKEKD 245  
DB 547 ----RHLPPTSSPWILQ-----SPVGDAIVDSEDD 573

RESULT 14  
ID NCRI HUMAN  
AC 075376; Q9UPV5; Q9UQ18; STANDARD; PRT; 2440 AA.  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Nuclear receptor co-repressor 1 (N-Cor1) (N-Cor).  
GN NCOR1 OR KIAA1047.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98393736; PubMed=9724795;  
RA Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.;  
RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-Cor/msin3/HDAC1 complex.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).  
RN [2]  
RP SEQUENCE OF 782-2440 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99397452; PubMed=10470851;  
RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";  
RL DNA Res. 6:197-205(1999).  
RN [3]  
RP SEQUENCE OF 974-2440 FROM N.A.  
RX MEDLINE=99375328; PubMed=1044436;  
RA Nagaya T., Chen K.-S., Fujieda M., Ohmori S., Richer J.K., Horwitz K.B., Lupski J.R., Seo H.;  
RT "Localization of the human nuclear receptor co-repressor (hn-Cor) gene between the CMT1A and the SMS critical regions of chromosome 17p11.2.";  
RL Genomics 59:339-341(1999).  
CC -!- FUNCTION: Mediates the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus preventing access of the basal transcription.  
CC -!- SUBUNIT: Interacts with HDAC7 (By similarity). Forms a large

corepressor complex that contains SIN3A/B and histone deacetylases HDAC1 and HDAC2. This complex associates with the thyroid (TR) and the retinoid acid receptors (RAR) in the absence of ligand.

-!- SUBCELLULAR LOCATION: Nuclear (By similarity).

CC DOMAIN: The N-terminal region contains repression functions that are divided into three independent repression domains (RD1, RD2 and RD3). The C-terminal region contains the nuclear receptor-interacting domains that are divided in two separate interaction domains (ID1 and ID2).

CC -!- DOMAIN: The two interaction domains (ID) contain a conserved sequence referred to as the cornr box. This motif is required and sufficient to permit binding to unliganded TR and RARs. Sequences flanking the CORNR BOX determine nuclear hormone receptor specificity.

CC -!- SIMILARITY: Contains 1 SANT-A domain.

CC -!- SIMILARITY: Contains 1 MYB-like domain.

CC -!- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF044209; AAC33550.1; -

DR EMBL; AB028970; BAA82999.1; -

DR EMBL; AB019524; BAA75814.1; -

DR TRANSFAC; T04687; -

DR Genew; HGNC:7672; NCOR1.

DR MIM; 600849; -

DR GO; GO:0003714; F:transcription co-repressor activity; TAS.

DR GO; GO:006366; P:transcription from Pol II promoter; TAS.

DR InterPro; IPR001005; MYB DNA binding.

DR Pfam; PF00249; MYB DNA-binding; 2.

DR SMART; SM00717; SANT; 2.

DR PROSITE; PS50090; MYB\_3; 1.

CC Nucleic protein; Transcription regulation; DNA-binding; Repressor; Coiled coil.

CC Coiled coil.

FT DOMAIN 174 216 COILED COIL (POTENTIAL).

FT DOMAIN 254 312 INTERACTION WITH SIN3A/B.

FT DOMAIN 299 328 COILED COIL (POTENTIAL).

FT DNA\_BIND 437 482 SANT-A (POTENTIAL).

FT DNA\_BIND 620 670 MYB.

FT DOMAIN 501 557 COILED COIL (POTENTIAL).

FT DOMAIN 607 617 PRO-RICH.

FT DOMAIN 988 1816 INTERACTION WITH ETO.

FT DOMAIN 2055 2059 CORNR BOX OF ID1.

FT DOMAIN 2267 2267 CORNR BOX OF ID2.

FT DOMAIN 58 64 POLY-GLN.

FT DOMAIN 593 603 POLY-ALA.

FT DOMAIN 1032 1035 POLY-PRO.

FT DOMAIN 1707 1712 POLY-ALA.

FT DOMAIN 1952 1963 POLY-SER.

FT CONFLICT 1014 1014 L -> V (IN REF. 2).

FT CONFLICT 1508 1509 PP -> SS (IN REF. 2).

FT CONFLICT 1561 1561 W -> R (IN REF. 2).

FT CONFLICT 1567 1567 Q -> H (IN REF. 2).

SQ SEQUENCE 2440 AA; 270263 MW; 60A4D7964D00EDAB CRC64;

Query Match 7.4%; Score 100; DB 1; Length 2440;

Best Local Similarity 20.7%; Pred. No. 10;

Matches 54; Conservative 37; Mismatches 82; Indels 88; Gaps 12;

QY 59 PSESTVSGNGIP-----EPQVYAPRPTDRLAVPPFAQRPRFRFQTPYQLQHEIDLPPT 114

DB 738 PENATSRGNTEPAVLEPTTAPTSPSLAV-----STKPAEDSVETQVN 785

QY 115 ISLS-----DGEE-----PPP--YQGFCTQLRDPPE-----QQLE 142

DB 786 DSAISAETAQMVDVQDQSAEGSVCDPPPPKADSDVDEVRVPPENHASKVEGDNTKRD 845

QY 143 LNRESVRAPPNRTIFDSLDMSARLGGPCP--PSSNSGISATCYGSGGRMGEP----- 194

DB 846 LDRASEKVEPR-----DEDLVVAQINARQPEFQSDNDSSATC-SADEVDVDEPERQRMFP 900

QY 195 ----PTYSEVIGHYPGSS-----FOHQOSSGPPSL-----LEGTRLH 227

DB 901 MDSKPSLLNPTGTSILVSSPLKPNPLDLPLQLHRAAVIPPMVSVCTPCNIPICTPVSGYALY 960

QY 228 HTHIAPLESAAINSEKDKQK 248

DB 961 QRHIKAMHESALLEEQRQRE 981

RESULT 15

EBN6\_EBV

ID EBN6\_EBV STANDARD; PRT; 992 AA.

AC P03204;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE EBN6-6 nuclear protein (EBNA-3C) (EBNA-4B).

GN BERP3-BERP4.

OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gamaherpesvirinae; Lymphocryptovirus.

OX NCBI\_TaxID=10377;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84270667; PubMed=6087149;

RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J., Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C., Tufnell P.S., Barrett B.G.;

RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";

RL Nature 310:207-211 (1984).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE=88155772; PubMed=2831394;

RA Petti L., Sample J., Wang F., Kieff E.;

RT "A fifth Epstein-Barr virus nuclear protein (EBNA3C) is expressed in latently infected growth-transformed lymphocytes.";

RL J. Virol. 62:1330-1338 (1988).

RN [3]

RP SUBCELLULAR LOCATION

RX MEDLINE=90266473; PubMed=2161150;

RA Petti L., Sample C., Kieff E.;

RT "Subnuclear localization and phosphorylation of Epstein-Barr virus latent infection nuclear proteins.";

RL Virology 176:563-574 (1990).

CC -!- FUNCTION: INVOLVED IN LATENT CYCLE.

CC -!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.

CC -!- SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, 5, AND 6.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; V01555; CAA24859.1; -

DR Pfam; PF05009; EBNA-3; 1.

KW Nuclear protein; Repeat.

FT DOMAIN 74 80 POLY-ARG.

FT DOMAIN 551 610 10 X 5 AA TANDEM REPEATS.

FT DOMAIN 741 779 3 X 13 AA TANDEM REPEATS.

SQ SEQUENCE 992 AA; 109129 MW; 39BEAB9BC515BD84 CRC64;

Query Match 7.4%; Score 99.5; DB 1; Length 992;

Best Local Similarity 23.9%; Pred. No. 4.1;

Matches 49; Conservative 16; Mismatches 83; Indels 57; Gaps 10;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:48:04 ; Search time 35 Seconds  
(without alignments)  
1857.979 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MAELFVQIIIVVMVMV.....PLESAAIWSKDKQKGHPL 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 23.1\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_todent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	95.1	237	4 Q8NER4	Q8ner4 homo sapien
2	820.5	60.9	292	11 Q8BYE2	Q8bye2 mus musculus
3	817.5	60.6	306	11 Q8BWJ4	Q8bwj4 mus musculus
4	662.5	49.1	200	11 Q8R5E2	Q8r5e2 mus musculus
5	252	18.7	53	4 Q8BYL8	Q8byl8 homo sapien
6	244	18.1	84	4 Q8NTR9	Q8ntr9 homo sapien
7	126	9.3	593	4 Q8IZL1	Q8izl1 homo sapien
8	125.5	9.3	494	4 Q86A18	Q86ai8 homo sapien
9	125.5	9.3	607	4 Q8IZ34	Q8iz34 homo sapien
10	125.5	9.3	634	4 Q8N6W3	Q8n6w3 homo sapien
11	125.5	9.3	634	4 Q75114	Q75114 homo sapien
12	123.5	9.2	496	10 Q8VZR8	Q8vzt8 arabidopsis
13	123.5	9.2	508	10 Q9M0H8	Q9m0h8 arabidopsis
14	123	9.1	462	11 Q8CF88	Q8cf88 mus musculus
15	122	9.1	357	10 Q8L7D7	Q8l7d7 arabidopsis
16	122	9.1	530	10 Q9LYK5	Q9lyk5 arabidopsis

17	118	8.8	291	6 Q97928	Q97928 pongo pygma
18	118	8.8	593	4 Q9H801	Q9h801 homo sapien
19	117.5	8.7	286	6 Q97957	Q97957 macaca fasc
20	117.5	8.7	301	6 Q97923	Q97923 hylobates l
21	115	8.5	299	6 Q97922	Q97922 gorilla gor
22	115	8.5	471	4 Q8N684	Q8n684 homo sapien
23	113.5	8.4	729	5 Q18660	Q18660 drosophila
24	112	8.3	471	11 Q8BTV2	Q8btv2 mus musculus
25	111.5	8.3	1182	4 Q99495	Q99495 homo sapien
26	111	8.2	727	10 Q9XIL9	Q9xil9 arabidopsis
27	110.5	8.2	300	10 Q8LIW2	Q8liw2 oryza sativ
28	110.5	8.2	778	4 Q8N5B2	Q8n5b2 homo sapien
29	109.5	8.1	596	5 Q8MZ04	Q8mz04 drosophila
30	109.5	8.1	610	5 Q8IPK6	Q8ipk6 drosophila
31	109.5	8.1	729	5 Q9VMS4	Q9vms4 drosophila
32	109.5	8.1	1099	5 Q24096	Q24096 drosophila
33	109.5	8.1	1099	5 Q24590	Q24590 drosophila
34	109.5	8.1	1105	5 Q9VA38	Q9va38 drosophila
35	109	8.1	1190	4 Q99621	Q99621 homo sapien
36	108.5	8.0	994	5 Q9BLN0	Q9bln0 theileria a
37	108.5	8.0	1167	5 Q9Y067	Q9y067 theileria a
38	108.5	8.0	3469	5 Q9U412	Q9u412 drosophila
39	108.5	8.0	3604	5 Q9VYK0	Q9vyk0 drosophila
40	107.5	8.0	1118	5 Q9VR13	Q9vr13 drosophila
41	107	7.9	313	6 Q97927	Q97927 pan paniscu
42	107	7.9	809	5 Q9NAN8	Q9nan8 caenorhabdi
43	107	7.9	2357	5 Q9U1M8	Q9u1m8 dictyosteli
44	106.5	7.9	358	10 Q9MAH5	Q9mah5 arabidopsis
45	106.5	7.9	939	10 Q8LJ66	Q8lj66 oryza sativ

## ALIGNMENTS

## RESULT 1

Q8NER4 PRELIMINARY; PRT; 237 AA.  
AC Q8NER4;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
DE PMPAL variant A protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Brunschwig E.B., Wilson K., Mack D., Dawson D., Lawrence E.,  
RA Willson J.K.V., Lu S., Nograti A., Swinler S., Beard L.,  
RA Lutterbaugh J.D., Willis J., Platzer P., Markowitz S.,  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY128643; AAM89277.1; --  
SQ SEQUENCE 237 AA; 26201 MW; A44A274EABFD930 CRC64;

Query Match 95.1%; Score 1282; DB 4; Length 237;  
Best Local Similarity 100.0%; Pred. No. 4.9e-117;  
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	16	MMWVVVITCLLSHYKLSARSFIRHSQGRRRDALSSGCLWPSSTVSGNGIPEQVY 75		
Db	1	MMWVVVITCLLSHYKLSARSFIRHSQGRRRDALSSGCLWPSSTVSGNGIPEQVY 60		
Qy	76	APPPPTDLAVPPFAQRERFHFQTPYLOHEIDLPTTISLSDGEEPPVGGPCTQLR 135		
Db	61	APPPPTDLAVPPFAQRERFHFQTPYLOHEIDLPTTISLSDGEEPPVGGPCTQLR 120		
Qy	136	DPEQQLNRESVRAPNRTTFDSDLMDSARLGCPSPSSNSGISATCYGSGRMEGPPP 195		
Db	121	DPEQQLNRESVRAPNRTTFDSDLMDSARLGCPSPSSNSGISATCYGSGRMEGPPP 180		
Qy	196	TYSVIGHYPOSSFOHQSSGPPSLLETRLUHHTHIAPLESAAIWSKDKQKGHPL 252		





```
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE DJ1059L7.1.2 (Androgen induced type 1b transmembrane protein (PMEP1),
DE isoform 2) (Fragment).
GN TMEP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121913; CAC32857.1; -.
KW Transmembrane.
FT NON_TER 53
SQ SEQUENCE 53 AA; 6064 MW; A3B563FB81F69782 CRC64;

Query Match          18.7%; Score 252; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.6e-17;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELFEVQIIIVVMVMVVITCLLGHYKLSARSFISRHSGRRREDALSS 53
Db 1 MAELFEVQIIIVVMVMVVITCLLGHYKLSARSFISRHSGRRREDALSS 53

RESULT 6
QNTN9 PRELIMINARY; PRT; 84 AA.
AC QNTN9;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE DJ1059L7.1.1 (Androgen induced type 1b transmembrane protein (PMEP1),
DE isoform 1) (Fragment).
GN TMEP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL121913; CAB88144.1; -.
KW Transmembrane.
FT NON_TER 84
SQ SEQUENCE 84 AA; 9236 MW; 7FCFCF1FFBEPBC92 CRC64;

Query Match          18.1%; Score 244; DB 4; Length 84;
Best Local Similarity 96.2%; Pred. No. 2.8e-16;
Matches 51; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAELFEVQIIIVVMVMVVITCLLGHYKLSARSFISRHSGRRREDALSS 53
Db 32 ITELEFVQIIIVVMVMVVITCLLGHYKLSARSFISRHSGRRREDALSS 84

RESULT 7
Q81ZL1 PRELIMINARY; PRT; 593 AA.
AC Q81ZL1;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Mucoepidermoid susceptibility protein.
GN MECT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Tonon G., Modi S., Kubo A., Coxon A., Kirsch I., Kaye F.J.;
RT "Exon1 swapping between MAML2 and MECT1 suggests a role for aberrant
RT NOTCH signaling in the genesis of mucoepidermoid tumors.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY040323; AAK93832.1; -.
SQ SEQUENCE 593 AA; 62529 MW; B7D8709B0C731D29 CRC64;

Query Match          9.3%; Score 126; DB 4; Length 593;
Best Local Similarity 24.6%; Pred. No. 0.001;
Matches 68; Conservative 30; Mismatches 80; Indels 98; Gaps 13;

QY 38 ISRHSQRRRE-----DALSSGCLWPSSESTVSGNGIPQVYVAPRP 80
Db 314 LSLSTEARRQASPLSLSPITQAVAMDALSLEQLPYAFFTQAGSQPQPQPQP 373
QY 81 TDLAVPPFAQRERFHRFPQTPYLYQHEIDLP-----PTISLSDGEBPPP--YQGPCTL 132
Db 374 -----PPASQQ-----PPPPPPQAPVRLPPGGPLLPASLTRGPPPLAVTPSSL 421
QY 133 QLRDPEQ-----QLSLNRESVRAPNRTIFDSDLMDSARLGGPCPPSSNSG 178
Db 422 PQSPENPGQPSMGIDIASAPALQQYRTSAGSPANQS-----PTSPVSNQG 467
QY 179 ISATCYSGGRME-----GPPPTYSEVIGHYPGSSFOHQSS-----GPSSLLEG 223
Db 468 FSP---GSSPQLEQFNMMENAISSSSLYS-----PGSTLNYQAAMMLTGTSGHSLPDS 518
QY 224 TRL---HHTHIAPL-----ESAAIWSKEKDKQKH 250
Db 519 QOLGYASHSGIPNIILTVTGESPPSLSKELTSHRGH 554

RESULT 8
Q96AI8 PRELIMINARY; PRT; 494 AA.
AC Q96AI8;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Colon;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017075; AAH17075.1; -.
DR InterPro; IPR002965; P rich extensn.
DR PRINTS; PR01217; PRICHEXTENSN.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 494 AA; 51628 MW; 720F127A9330ECAP CRC64;

Query Match          9.3%; Score 125.5; DB 4; Length 494;
Best Local Similarity 24.7%; Pred. No. 0.0093;
Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;

QY 38 ISRHSQRRRE-----DALSSGCLWPSSESTVSGNGIPQVYVAPRP 80
Db 174 LSLSTEARRQASPLSLSPITQAVAMDALSLEQLPYAFFTQAGSQPQPQPQP 233
QY 81 TDLAVPPFAQRERFHRFPQTPYLYQHEIDLP-----PTISLSDGEBPPP--YQGPCTL 132
Db 234 -----PPASQQ-----PPPPPPQAPVRLPPGGPLLPASLTRGPPPLAVTPSSL 281
QY 133 QLRDPEQ-----QLSLNRESVRAPNRTIFDSDLMDSARLGGPCPPSSNSG 178
Db 282 PQSPENPGQPSMGIDIASAPALQQYRTSAGSPANQS-----PTSPVSNQG 327
```

QY 179 ISATCYGSGRMEGPPPTYSVEIGHYVPGSSFOHQ 213  
 Db 328 FS-----PGSSPQHTSTLGSVFGDAYVEQ 352  
 RESULT 9  
 Q81Z34 PRELIMINARY; PRT; 607 AA.  
 ID Q81Z34  
 AC Q81Z34  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC023614; AAH23614.1; --  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 607 AA; 64275 MW; E9380FE66B437CB5 CRC64;

Query Match 9.3%; Score 125.5; DB 4; Length 607;  
 Best Local Similarity 24.7%; Pred. No. 0.0012;  
 Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;  
 QY 38 ISRHSQGRRE-----DALSSGCLWPSESTVSGNGIPEQVYAPRP 80  
 Db 287 LSLSTARQQASPTLSPLSLPITQAVAMDALSLEQLPYAFFTQAGSQQPPQP 346  
 QY 81 TDLAVPPFAQRERFHRFQPTYPYLQHEIDL-----PTISLSDGEPPPP--YQGPCTL 132  
 Db 347 -----PPASQ-----PPPPPPQAPVRLPPGGLPLPSASLTGPGPPPLAVTVPSL 394  
 QY 133 QLRDEQ-----QLELNRESVRAPNRTIFDSDLMSARLGCPGPPSSNSG 178  
 Db 395 PQSPENPGQSMGIDIASAPALQYRTSAGSPANQS-----PTSPVSNQ 440  
 QY 179 ISATCYGSGRMEGPPPTYSVEIGHYVPGSSFOHQ 213  
 Db 441 FS-----PGSSPQHTSTLGSVFGDAYVEQ 465

RESULT 10  
 Q8N6W3 PRELIMINARY; PRT; 634 AA.  
 ID Q8N6W3  
 AC Q8N6W3  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-NAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Similar to KIAA0616 protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC028050; AAH28050.1; --  
 DR InterPro; IPR002965; P rich exten.  
 DR PRINTS; PR01217; PRICHEXTEN.  
 SQ SEQUENCE 634 AA; 67269 MW; ECD53DD389D041DE CRC64;  
 Query Match 9.3%; Score 125.5; DB 4; Length 634;  
 Best Local Similarity 24.7%; Pred. No. 0.0013;  
 Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;  
 QY 38 ISRHSQGRRE-----DALSSGCLWPSESTVSGNGIPEQVYAPRP 80  
 Db 314 LSLSTARQQASPTLSPLSLPITQAVAMDALSLEQLPYAFFTQAGSQQPPQP 373  
 QY 81 TDLAVPPFAQRERFHRFQPTYPYLQHEIDL-----PTISLSDGEPPPP--YQGPCTL 132  
 Db 374 -----PPASQ-----PPPPPPQAPVRLPPGGLPLPSASLTGPGPPPLAVTVPSL 421

Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;  
 QY 38 ISRHSQGRRE-----DALSSGCLWPSESTVSGNGIPEQVYAPRP 80  
 Db 314 LSLSTARQQASPTLSPLSLPITQAVAMDALSLEQLPYAFFTQAGSQQPPQP 373  
 QY 81 TDLAVPPFAQRERFHRFQPTYPYLQHEIDL-----PTISLSDGEPPPP--YQGPCTL 132  
 Db 374 -----PPASQ-----PPPPPPQAPVRLPPGGLPLPSASLTGPGPPPLAVTVPSL 421  
 QY 133 QLRDEQ-----QLELNRESVRAPNRTIFDSDLMSARLGCPGPPSSNSG 178  
 Db 422 PQSPENPGQSMGIDIASAPALQYRTSAGSPANQS-----PTSPVSNQ 467  
 QY 179 ISATCYGSGRMEGPPPTYSVEIGHYVPGSSFOHQ 213  
 Db 468 FS-----PGSSPQHTSTLGSVFGDAYVEQ 492  
 RESULT 11  
 O75114 PRELIMINARY; PRT; 634 AA.  
 ID O75114  
 AC O75114  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Hypothetical protein KIAA0616 (Fragment).  
 GN KIAA0616.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98403880; PubMed=9734811;  
 RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,  
 Nomura N., Ohara O.;  
 RA "Prediction of the coding sequences of unidentified human genes. X.  
 The complete sequences of 100 new cDNA clones from brain which can  
 code for large proteins in vitro."  
 RL DNA Res. 5:169-176(1998).  
 RN [2]  
 RP SEQUENCE OF 27-440 FROM N.A.  
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,  
 Burkhardt-Schultz K., Gordon L., Dias J., Kyle A., Brower A.,  
 RA Stilwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,  
 RA Ganes J., Danganan L., Erler A., Christensen M., Georgescu A.,  
 RA Avila J., Liu S., Attix C., Andreise T., Trankheim M.,  
 RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,  
 RA Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,  
 RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;  
 RA "Sequence analysis of a 1.9 Mb contig in 19p12 between UBA52 and  
 D19S455."  
 RT D19S455;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB014516; BAA31591.1; --  
 DR EMBL; AC006123; AAC97072.1; --  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 634 AA; 66992 MW; C2D3BE373C0979A1 CRC64;

Query Match 9.3%; Score 125.5; DB 4; Length 634;  
 Best Local Similarity 24.7%; Pred. No. 0.0013;  
 Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;  
 QY 38 ISRHSQGRRE-----DALSSGCLWPSESTVSGNGIPEQVYAPRP 80  
 Db 314 LSLSTARQQASPTLSPLSLPITQAVAMDALSLEQLPYAFFTQAGSQQPPQP 373  
 QY 81 TDLAVPPFAQRERFHRFQPTYPYLQHEIDL-----PTISLSDGEPPPP--YQGPCTL 132  
 Db 374 -----PPASQ-----PPPPPPQAPVRLPPGGLPLPSASLTGPGPPPLAVTVPSL 421

```
QY 133 QLRDEQ-----QLELNRESVRAPPNRTIFDSDLMDLSARLGGCPSPSSNSG 178
Db 422 POSPPNFCQPSMGIDIASAPALQYRTSAGSPANQS-----PTSPVSNQG 467
QY 179 ISATCYGSGRMEGPPPTPYSEVIGHYVPGSSFOHQ 213
Db 468 FS-----PGSSPQHTSLGVSFGDAYVEQ 492

RESULT 12
Q8VZR8
ID Q8VZR8 PRELIMINARY; PRT; 496 AA.
AC Q8VZR8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative proline-rich protein.
GN AT4G28300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Jones T., Carninci P., Chen H., Cheuk R.,
RA Havaahizaki Y., Iehida J., Bowser L., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesena E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene At4g28300 (GI:15235249).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Havaahizaki Y., Iehida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY063895; AAL36251.1; -
DR EMBL; AY096387; AAM20028.1; -
DR InterPro; IPR002965; P rich extensn.
DR PRINTS; PR01217; PRICHEXTENS.
SQ SEQUENCE 496 AA; 54106 MW; 1BF54098B9A79429 CRC64;

Query Match 9.2%; Score 123.5; DB 10; Length 496;
Best Local Similarity 26.4%; Pred. No. 0.0015;
Matches 51; Conservative 18; Mismatches 71; Indels 53; Gaps 10;

QY 70 PEPQVYAPRPTRDLAVPP-----FAQREFRHFQTPY-----YLQHEIDLPTISLS 118
Db 259 PPSQLQAPPAQSQFMPPPAPSHPSAQTQSPQYQQNWPPQOPARPOSSGGYPTYSAP 318
QY 119 DGEPPP-----PYQGCTQLRDPQQLNRESVRAPPNRTIFDSDLMDLSAR 166
Db 319 PGNQPPVESLPSSMQMSPYSGP-----PQSMQAYGYGAAPPQAPPQTKMSYSPQ 371
QY 167 LG-----GPCPPSSNSGISATCYSGGRMEGPP--PYSEVIGHY-----PGSSFOHQ 212
Db 372 TGDGYLPSPGPPPP---SGYANAMY-EGGRMYPFPQPOQQQQAHYLQGPQGGYSPQPH 427
QY 213 QSSG-----PPSL 220
Db 428 QAGGNIGAPPVL 440

Query Match 9.2%; Score 123.5; DB 10; Length 496;
Best Local Similarity 26.4%; Pred. No. 0.0015;
Matches 51; Conservative 18; Mismatches 71; Indels 53; Gaps 10;

QY 70 PEPQVYAPRPTRDLAVPP-----FAQREFRHFQTPY-----YLQHEIDLPTISLS 118
Db 271 PPSQLQAPPAQSQFMPPPAPSHPSAQTQSPQYQQNWPPQOPARPOSSGGYPTYSAP 330
QY 119 DGEPPP-----PYQGCTQLRDPQQLNRESVRAPPNRTIFDSDLMDLSAR 166
Db 331 PGNQPPVESLPSSMQMSPYSGP-----PQSMQAYGYGAAPPQAPPQTKMSYSPQ 383
QY 167 LG-----GPCPPSSNSGISATCYSGGRMEGPP--PYSEVIGHY-----PGSSFOHQ 212
Db 384 TGDGYLPSPGPPPP---SGYANAMY-EGGRMYPFPQPOQQQQAHYLQGPQGGYSPQPH 439
QY 213 QSSG-----PPSL 220
Db 440 QAGGNIGAPPVL 452

RESULT 14
Q8CF88
ID Q8CF88 PRELIMINARY; PRT; 462 AA.
AC Q8CF88;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to hypothetical protein FLJ12529.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Straussberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038812; AAH38812.1; -
KW Hypothetical protein.
SQ SEQUENCE 462 AA; 51056 MW; 42D511ECC74E228E CRC64;

Query Match 9.1%; Score 123; DB 11; Length 462;
Best Local Similarity 25.4%; Pred. No. 0.0015;
```

Matches 50; Conservative 17; Mismatches 70; Indels 60; Gaps 8;

QY 27 LSHYKLSARSPFISRSQGRREDALSSGCLWPSSTVSGN----- 67  
Db 165 LSQFAQAKRIIPRAHSRSDSD--SADGRATSENLVSSARVDKPPSVLVYFNRPSPA 222  
QY 68 ----GIPEQVYAPRPTDRLAVPPFAQREHFRFOPTYPYLQHEIDLPTTISLSDGEEP 123  
Db 223 LPLMGLPPPIPPPPPLSSFGVPPPPGHIQHLMPPPP-----RLPPHLAV-----P 271  
QY 124 PYYQGPCTLQLRDPEQQLNRESVRAPNRTIFDSDLMSARLGCPGPPSSNSGISATC 183  
Db 272 PPGAIPPALHL-----NPAFFPPPNATV-----GP-PDITYMKASTPY 308  
QY 184 YGSGGRMEGPP-tyse 199  
Db 309 NHGSRDGGPPSTVSE 325

## RESULT 15

Q8L7D7 PRELIMINARY; PRT; 357 AA.  
AC Q8L7D7;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN At5G14540.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,  
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,  
RA Theologis A., Davis R.W.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY136314; RAM96980.1; -  
DR InterPro: IPR002965; P-rich\_extensn.  
DR InterPro: IPR000449; UBA domain.  
DR PRINTS: PR01217; PRICHEXTENS.  
KW Hypothetical protein.  
SQ SEQUENCE 357 AA; 38339 MW; 499FDF163DDF8030.CRC64;

Query Match 9.1%; Score 122; DB 10; Length 357;  
Best Local Similarity 29.1%; Pred. NO. 0.0014;  
Matches 48; Conservative 8; Mismatches 71; Indels 38; Gaps 8;

QY 65 SGNIGPEQV---YAPRPTDRLAVPPFAQREHFRFOPTYPYLQHEIDLPTTISLSDGE 121  
Db 111 SQSQPPPTIQPYQPPPTQSLHQPYPQPQ----QPQYDQ-----QPPQLQHPSGY 161  
QY 122 EP--PPYQGPCTLQLRDPEQQLNRESVRAPNRTIFDSDLMSARLGCPGPPS----- 174  
Db 162 NPEEPY-----PQSYPPN--PPRQPPSPGSAFSQQYNNAPTTPSPMYDGP 209  
QY 175 ----SNSGISATCYGSGGRMEGPPPTYSVIGHYPGSSFOHQSSG 216  
Db 210 GGRSNSGPPSGSPSYPTGPPSQY----GNTPSVKPTHQSGSG 250

Search completed: December 4, 2003, 19:51:32  
Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:44:19 ; Search time 41 Seconds  
(without alignments)  
975.588 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MAELFVQIIIVVMVMV.....PLESAIWSKQKQKHPL (252)

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	252	21 AAB18449	A human TANGO 261
2	1348	100.0	252	21 AAB01388	Neuron-associated
3	1348	100.0	252	23 ABG61805	Prostate cancer-as
4	1348	100.0	285	24 ABP75660	Human secretory po
5	1345	99.8	252	21 AAB18461	A human TANGO 261
6	1345	99.8	252	21 AAB18462	A human TANGO 261
7	1345	99.8	252	21 AAB18463	A human TANGO 261
8	1340	99.4	287	23 AAU78231	Human Mechanically
9	1282	95.1	237	22 AAM38712	Human polypeptide

10	1172.5	87.0	274	23	AAU78236	Mouse Mechanically
11	1133	84.1	207	22	AAM40498	Human polypeptide
12	1066.5	79.1	217	21	AAB18450	A murine TANGO 261
13	1065.5	79.0	241	23	ABG77052	Prostate specific
14	1063.5	78.9	217	21	AAB18464	A murine TANGO 261
15	1063.5	78.9	217	21	AAB18465	A murine TANGO 261
16	1063.5	78.9	217	21	AAB18466	A murine TANGO 261
17	844.5	62.6	269	21	AAU73422	Human secreted pro
18	843.5	62.6	306	19	AAW60673	Manic-depressive i
19	842.5	62.5	334	22	ABG13117	Novel human diagno
20	796	59.1	261	22	ABG20316	Novel human diagno
21	792.5	58.8	288	19	ABW60674	Manic-depressive i
22	791.5	58.7	316	22	ABG13118	Novel human diagno
23	765.5	56.8	235	22	ABG13123	Novel human diagno
24	672	49.9	270	20	AAU74136	Human prostate tum
25	600	44.5	204	22	ABG13125	Novel human diagno
26	566.5	42.0	475	22	ABG28310	Novel human diagno
27	566.5	42.0	485	22	ABG13126	Novel human diagno
28	126	9.3	593	24	AAU79911	MECT1. Homo sapie
29	125.5	9.3	650	22	AAU28167	Novel human secret
30	125.5	9.3	657	22	AAU28355	Novel human secret
31	122	9.1	522	21	AAG13630	Arabidopsis thalia
32	122	9.1	539	21	AAG13629	Arabidopsis thalia
33	122	9.1	540	21	AAG13628	Arabidopsis thalia
34	118	8.8	593	22	AAB95646	Human protein sequ
35	118	8.8	804	22	ABG20282	Novel human diagno
36	116	8.6	418	23	ABG66738	Human novel polype
37	115	8.5	471	24	AAO19620	Human nucleic acid
38	114	8.5	1518	24	ABJ18375	Breast specific re
39	113.5	8.4	260	21	AGG42023	Arabidopsis thalia
40	113.5	8.4	395	21	AAG42022	Arabidopsis thalia
41	112.5	8.3	532	22	AAU33219	Novel human secret
42	111.5	8.3	1185	20	AAU33497	Human atrophin I p
43	111	8.2	1401	22	ABG24037	Novel human diagno
44	110.5	8.2	778	17	AAR94900	Human tastin. Hom
45	110.5	8.2	778	21	AAB23359	Human tastin prote

ALIGNMENTS

RESULT 1  
AAB18449  
ID AAB18449 standard; Protein; 252 AA.  
XX AC AAB18449;  
XX AC AAB18449;  
XX AC AAB18449;  
DT 15-JAN-2001 (first entry)  
DE A human TANGO 261 polypeptide.

TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
cellular proliferation; cellular differentiation; cancer;  
von Willebrand factor-associated disorder; cell trafficking; cancer;  
hematopoietic associated disease; atelectasis; pulmonary congestion;  
oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
intestinal disorder; spleen associated disease; renal disorder;  
cardiovascular disorder; ischemic heart disease; hydrocephalus;  
brain herniation; istrogenic disease; inflammation; meningitis;  
Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..28  
FT Protein /note= "signal peptide"  
FT Protein 29..253  
XX /note= "mature protein"  
XX WO200052022-A1.  
XX 08-SEP-2000.  
PD

XX 01-MAR-2000; 2000WO-US05226.  
PF 01-MAR-1999; 99US-0122458.  
PR (MILL-) MILLENNIUM PHARM INC.  
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX WPI; 2000-579269/54.  
DR N-PSDB; AAA75151.  
XX Novel human and murine secreted proteins designated TANGO 216, 261,  
PT 262, 266 and 267 useful as modulating agents of cellular processes,  
PT e.g. for treating cancer -  
XX Claim 2; Fig 5; 175pp; English.  
XX The present sequence represents a human TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate  
CC cellular adhesion. The proteins can be used to treat any von Willebrand  
CC factor-associated disorder, regulate extracellular matrix structuring,  
CC cellular adhesion, and cell trafficking and/or migration, modulate  
CC cellular interactions, modulate cell adhesion in proliferative  
CC disorders, such as cancer, modulate the proliferation, differentiation,  
CC and/or function of cells that appear in the bone marrow, and leukocytes,  
CC treat bone marrow, blood and hematopoietic associated diseases and  
CC disorders, atelectasis, pulmonary congestion or oedema, emphysema,  
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
CC disorders, spleen associated diseases, modulate renal disorders, treat  
CC cardiovascular disorders such as ischemic heart disease, modulate the  
CC proliferation, differentiation, and/or function of bone and cartilage  
CC cells and to treat bone and/or cartilage associated diseases or  
CC disorder. They may also be used to treat disorders associated with the  
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic  
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,  
CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
CC disorders.  
XX Sequence 252 AA;  
SQ  
Query Match 100.0%; Score 1348; DB 21; Length 252;  
Best Local Similarity 100.0%; Pred. No. 1.3e-111;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAELEFVQIIIVVVMVWVITCLSHYKLSARSFISRSQGRREDALSSGCLWPS 60  
Db 1 MAELEFVQIIIVVVMVWVITCLSHYKLSARSFISRSQGRREDALSSGCLWPS 60  
QY 61 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQRFRHPTQYTPYLOHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQRFRHPTQYTPYLOHEIDLPTTISLSDG 120  
QY 121 EEPYQCGCTQLDPPQOELNRESVRAPNRTIFDSDLMSARLGGPCPPSSNSGIS 180  
Db 121 EEPYQCGCTQLDPPQOELNRESVRAPNRTIFDSDLMSARLGGPCPPSSNSGIS 180  
QY 181 ATCYGSGRMGEPPTTYSEVIGHYFGSGFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240  
Db 181 ATCYGSGRMGEPPTTYSEVIGHYFGSGFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240  
QY 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252  
RESULT 2  
AAB01388  
ID AAB01388 standard; Protein; 252 AA.  
XX

AC AAB01388;  
XX 20-OCT-2000 (first entry)  
XX Neuron-associated protein.  
DE  
XX Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KW Alzheimer's disease; Pick's disease; Huntington's disease;  
KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
KW peripheral nervous system; PNS; myopathy; schizophrenia;  
KW acinic keratosis; arteriosclerosis; atherosclerosis; bursitis;  
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KW AIDS; Addison's disease; adult respiratory distress syndrome;  
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KW Werner syndrome; trauma; human.  
XX Homo sapiens.  
OS  
XX Location/Qualifiers  
Key  
FT Peptide 1..33  
FT Region 9..27  
FT /label= Signal peptide  
FT /label= Transmembrane region  
FT Modified-site 33  
FT /note= "Potential phosphorylation site"  
FT Modified-site 65..68  
FT /note= "Glycosaminoglycan attachment site"  
FT Modified-site 81  
FT /note= "Potential phosphorylation site"  
FT Modified-site 116  
FT /note= "Potential phosphorylation site"  
FT Modified-site 118  
FT /note= "Potential phosphorylation site"  
FT Modified-site 147  
FT /note= "Potential phosphorylation site"  
FT Modified-site 153  
FT /note= "Potential phosphorylation site"  
FT Modified-site 155  
FT /note= "Potential glycosylation site"  
FT Modified-site 164  
FT /note= "Potential phosphorylation site"  
FT Modified-site 196  
FT /note= "Potential phosphorylation site"  
FT Modified-site 219  
FT /note= "Potential phosphorylation site"  
XX WO200034477-A2.  
XX 15-JUN-2000.  
XX 10-DEC-1999; 99WO-US30408.  
XX 11-DEC-1998; 98US-0210083.  
XX 11-DEC-1998; 98US-9123456.  
XX 09-FEB-1999; 99US-0119365.  
XX 16-MAR-1999; 99US-0124687.  
XX (INCY-) INCYTE PHARM INC.  
XX Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;  
PI Lu DAM, Azimzai Y;  
XX WPI; 2000-423423/36.  
XX N-PSDB; AAA47429.  
XX New human neuron-associated proteins and polynucleotides encoding them,  
PT useful for diagnosis, treatment and prevention of cell proliferative  
PT disorders including cancer, neuronal and neurological disorders

self-

XX Claim 1; Page 113-114; 145pp; English.

XX Human neuron-associated proteins (NEUAP) can be used for for

CC treating or preventing a disorder associated with decreased

CC expression or activity of NEUAP. Antagonists of NEUAP are useful for

CC treating or preventing disorder associated with increased expression

CC or activity of NEUAP. NEUAP or their fragments or derivatives are

CC useful for treating neurological disorder such as epilepsy, ischemic

CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's

CC disease, Pick's disease, Huntington's disease, dementia and

CC Parkinson's disease. NEUAPs are also useful for treating other

CC demyelinating diseases, bacterial and viral meningitis, prion

CC diseases including kuru, Creutzfeldt-Jacob disease, nutritional and

CC metabolic diseases of the nervous system, neurofibromatosis, other

CC developmental disorders of the central nervous system, cerebral

CC palsy, neuromuscular disorders, autonomic nervous system disorders,

CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and

CC other neuromuscular disorders, peripheral nervous system disorders,

CC inherited, metabolic, endocrine, and toxic myopathies, mental

CC disorders including mood, anxiety and schizophrenic disorders, a cell

CC proliferative disorder such as actinic keratosis, arteriosclerosis,

CC atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective

CC tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal

CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,

CC bone marrow, brain, breast, cervix, and an autoimmune/inflammatory

CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's

CC disease, adult respiratory distress syndrome, allergies, ankylosing

CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,

CC complications of cancer, hemodialysis, and extracorporeal circulation,

CC viral, bacterial, fungal parasitic, protozoal, and helminthic

CC infections, and trauma. This protein was given the incyte ID no.

XX 1871288CD1.

SQ Sequence 252 AA;

Query Match 100.0%; Score 1348; DB 21; Length 252;

Best Local Similarity 100.0%; Pred. No. 1.3e-111;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRSQGRRRDALSSGCLWPS 60

DB 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRSQGRRRDALSSGCLWPS 60

QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

DB 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

QY 121 BEPPPYQGCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGGPCPPSSNSGIS 180

DB 121 BEPPPYQGCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGGPCPPSSNSGIS 180

QY 181 ATCVSGGGRMEGPPPTYSEVIGHYPGSGSFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240

DB 181 ATCVSGGGRMEGPPPTYSEVIGHYPGSGSFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240

QY 241 SKEKDQKQGHPL 252

DB 241 SKEKDQKQGHPL 252

RESULT 3

ID ABG61805

XX ABG61805 standard; Protein; 252 AA.

XX AC ABG61805;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #6.

XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX

OS Mammalia.

XX WO200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

PR 08-DEC-2000; 2000US-0733288.

PR 08-DEC-2000; 2000US-0733742.

PR 24-JAN-2001; 2001US-263957P.

PR 16-MAR-2001; 2001US-276791P.

PR 16-MAR-2001; 2001US-276888P.

PR 06-APR-2001; 2001US-281922P.

PR 24-APR-2001; 2001US-286214P.

PR 30-APR-2001; 2001US-0847046.

PR 04-MAY-2001; 2001US-288589P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

PI WPI; 2002-471335/50.

XX N-PSDB; ABK92120.

DR Detecting a prostate cancer-associated transcript in a cell in a

PT patient, useful for diagnosing prostate cancer (PC) or screening

PT modulators of PC, by determining if prostate cancer-associated genes

PT are expressed in a prostate tissue -

XX Claim 27; Page 305; 436pp; English.

XX The present invention relates to methods of detecting a prostate

CC cancer-associated transcript in a cell from a patient. The method

CC comprises contacting a biological sample from the patient with

CC prostate cancer-associated polynucleotides (designated PC genes) that

CC selectively hybridise to a sequence that is at least 80% identical

CC to them. The prostate cancer-associated polynucleotide sequences

CC are differentially expressed in prostate tumour tissue or in

CC prostate cancer and are derived from the tissues of various

CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).

CC The methods of the invention are useful for diagnosing and treating

CC prostate cancer in mammals. The prostate cancer-associated genes are

CC useful for diagnosing or treating prostate cancer, as well as for

CC identifying modulators of prostate cancer or agents that inhibit

CC prostate cancer. The nucleic acid sequences are particularly useful

CC in gene therapy, as a vaccine or in antisense applications.

CC ABG61800-ABG61944 represent prostate cancer-associated proteins.

XX Sequence 252 AA;

Query Match 100.0%; Score 1348; DB 23; Length 252;

Best Local Similarity 100.0%; Pred. No. 1.3e-111;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRSQGRRRDALSSGCLWPS 60

DB 1 MAELFVQIIIVVMVMVVITCLSHYKLSARSFISRSQGRRRDALSSGCLWPS 60

QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

DB 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

QY 121 BEPPPYQGCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGGPCPPSSNSGIS 180

DB 121 BEPPPYQGCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGGPCPPSSNSGIS 180

QY 181 ATCVSGGGRMEGPPPTYSEVIGHYPGSGSFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240

DB 181 ATCVSGGGRMEGPPPTYSEVIGHYPGSGSFQHQSSGPPSLLEGLRHHHTHAPLESAAIW 240

QY 241 SKEKDQKQGHPL 252

Db 241 SKKDKQKGHPL 252  
|||||

## RESULT 4

ABP75660  
ID ABP75660 standard; Protein; 285 AA.

XX AC ABP75660;

XX DT 10-FEB-2003 (first entry)

XX DE Human secretory polypeptide SPTM SEQ ID NO 844.

XX KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
XX asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
KW anti-inflammatory; immunosuppressive; neuroprotective; nontropic;  
KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;  
KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
KW secretory polynucleotide; secretory protein.

XX OS Homo sapiens.

XX PN WO200283876-A2.

XX PD 24-OCT-2002.

XX PF 27-MAR-2002; 2002WO-US09921.

XX PR 29-MAR-2001; 2001US-280067P.

XX PR 29-MAR-2001; 2001US-280068P.

XX PR 16-MAY-2001; 2001US-291280P.

XX PR 17-MAY-2001; 2001US-291829P.

XX PR 19-JUN-2001; 2001US-291849P.

XX PR 20-JUN-2001; 2001US-299776P.

XX PR 20-JUN-2001; 2001US-300001P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX PI Daffour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amesey SR;

XX PI Daugherty SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

CC breast, cervix or prostate). The present sequence is one of the SPTM  
CC proteins of the invention (ABP75384-ABP75962).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 285 AA;

Query Match 100.0%; Score 1348; DB 24; Length 285;

Best Local Similarity 100.0%; Pred. No. 1.5e-111;

Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRDLASSEGCLWPS 60

Db 34 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRDLASSEGCLWPS 93

Qy 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERPHRFPQTPYVLQHEIDLPPTISLDG 120

Db 94 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERPHRFPQTPYVLQHEIDLPPTISLDG 153

Qy 121 BEPPPYQPCPTQLRDPEQQLNRESVRAPNPNTIPSDMLDSARLGGCPPSSNSGIS 180

Db 154 BEPPPYQPCPTQLRDPEQQLNRESVRAPNPNTIPSDMLDSARLGGCPPSSNSGIS 213

Qy 181 ATCYGSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSPPSLLEGTRLHHTHIAPLESAI 240

Db 214 ATCYGSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSPPSLLEGTRLHHTHIAPLESAI 273

Qy 241 SKKDKQKGHPL 252

Db 274 SKKDKQKGHPL 285

## RESULT 5

AAB18461

ID AAB18461 standard; Protein; 252 AA.

XX AC AAB18461;

XX DT 15-JAN-2001 (first entry)

XX DE A human TANGO 261 polypeptide clone.

XX KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;

XX KW cellular proliferation; cellular differentiation; cellular adhesion;

XX KW von Willebrand factor-associated disorder; cell trafficking; cancer;

XX KW hematopoietic associated disease; atelectasis; pulmonary congestion;

XX KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;

XX KW intestinal disorder; spleen associated disease; renal disorder;

XX KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

XX KW brain herniation; iatrogenic disease; inflammation; meningitis;

XX KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;

XX KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

XX OS Homo sapiens.

XX PN WO200052022-A1.

XX PD 08-SEP-2000.

XX PF 01-MAR-2000; 2000WO-US05226.

XX PR 01-MAR-1999; 99US-0122458.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX DR WPI; 2000-579269/54.

XX DR N-PSDB; AAA75163.

XX PT Novel human and murine secreted proteins designated TANGO 216, 261,

XX PT 262, 266 and 267 useful as modulating agents of cellular processes,

XX PS Claim 27; SEQ ID NO 844; 458pp + Sequence Listing; English.

XX CC The invention relates to a secretory polynucleotide (designated sptm)

XX CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a

XX CC naturally occurring polynucleotide sequence at least 90 % identical to

XX CC the polynucleotide sequence, a polynucleotide complementary to them, or an

XX CC RNA equivalent of them. The polypeptide or polynucleotide are useful for

XX CC treating, preventing or diagnosing a disease or condition associated with

XX CC the expression of functional SPTM. These are particularly useful for

XX CC diagnosing, treating or preventing autoimmune/inflammatory disorders

XX CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

XX CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,

XX CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

XX CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,

XX CC schizophrenia or amnesia), or cell proliferative disorders (e.g.

XX CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,

XX CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,



e.g. for treating cancer -

Disclosure; Page -; 175pp; English.

RAB18461-63 represent human TANGO 261 proteins. The specification also describes TANGO 286, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or disorder. They may also be used to treat disorders associated with the ovaries, and cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders. note: the present sequence does not appear in the specification; it was created using information provided.

Sequence 252 AA;

Query Match 99.8%; Score 1345; DB 21; Length 252;

Best Local Similarity 99.6%; Pred. No. 2.3e-111;

Matches 251; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	M A E L F E V Q I I I I V V M M V M V V I T C L L S H Y K L S A R S F I S R H S Q Q R R R E D A L S E G C L W P S	60
Db	1	M A D L F E V Q I I I I V V M M V M V V I T C L L S H Y K L S A R S F I S R H S Q Q R R R E D A L S E G C L W P S	60
Qy	61	E S T V S G N G I P E P O N V A P P R P T D R L A V P P F A Q R F R H R F Q P T Y P Y L Q E I D L P T T I S L S D G	120
Db	61	E S T V S G N G I P E P O N V A P P R P T D R L A V P P F A Q R F R H R F Q P T Y P Y L Q E I D L P T T I S L S D G	120
Qy	121	E B P P Y Q G P C T L Q L R D P P Q Q L E A N R E S V R A P P N T I F D S D L M D S A R L G G C P P S S N S G I S	180
Db	121	E B P P Y Q G P C T L Q L R D P P Q Q L E A N R E S V R A P P N T I F D S D L M D S A R L G G C P P S S N S G I S	180
Qy	181	A T C Y G S G R M E G P P T Y S E V I G H Y P G S S F Q H Q Q S G P P S L L E G T R L H T H T I A P L E S A A I W	240
Db	181	A T C Y G S G R M E G P P T Y S E V I G H Y P G S S F Q H Q Q S G P P S L L E G T R L H T H T I A P L E S A A I W	240

## RESULT 6

AAB18462

ID AAB18462 standard; Protein; 252 AA.

AC AAB18462;

DT 15-JAN-2001 (first entry)

DE A human TANGO 261 polypeptide clone.

TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

brain herniation; iatrogenic disease; inflammation; meningitis; Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease; multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

**Homo sapiens.**

WO200052022-A1.

08-SEP-2000.

01-MAR-2000: 2000WO-US05226.

01-MAR-1999: 99US-0122458.

(MILL-) MILLENNIUM PHARM INC.

Barnes TM, Holtzman DA, Sharp JD, Fraser CC:

WPI: 2000-579269/54.

WFI, 2000-373263/  
N-PSDB: AAA75164.

Novel human and murine secreted proteins designated TANGO 216, 261, 262, 266 and 267 useful as modulating agents of cellular processes, e.g. for treating cancer -

Disclosure: Page -: 175pp: English.

AAB18461-63 represent human TANGO 261 proteins. The specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or disorder. They may also be used to treat disorders associated with the ovaries, and cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders. note: the present sequence does not appear in the specification; it was created using information provided.

Sequence 252 AA;

99.8%: Score 1345: DB 21: Length 252:

Best Local Similarity 99.6%; Pred. No. 2.3e-111;

```
Matches 251; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

1 MAELEFVQIIIVVMMVMVVITCLSHYKLSARSFIRHSQRRREDALSSEGCLWPS 60

1 MAELEFVQIIIVVMVMVVITCLSHYKLSARFISRHSQRRRDDALSSEGCLWPS 60

61 ESTVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQTPYPLQHEIDLPTISLSDG 120

61 ESTVSGNGIPEPOVYAPPRPTDRLAVPPEAORERFHRFOPTYPYLOHEIDLPPTISLSDG 120

121 EEPPPYOGPCTLOLRDPEOOLELNRESVRAPPNRTIFDSDLMD SARLGPGCPPSSNSGIS 180

121 EEPYPYOGPCTTQI.RDPEOOI.EI.NRESVRAPPNRTTFDSI.MDSARI.GGPCPPSSNSGTS 180

181 ATCTGSGGRMEGPPPTVSEVTGHVPGSSFOHOOSSGPPPSI.I.EGTRI.HHTHTAPI.ESAATW 240

181 ATCTGCGCGPMEGPPPTYSFVICHYPCSSFHQOSSCPBSLIECTPIHHTHIIRIESAAIW 240

QY 241 SKEKDQKQHPL 252  
 DB 241 SKEKDQKQHPL 252

RESULT 7  
 AAB18463  
 ID AAB18463 standard; Protein; 252 AA.  
 XX  
 AC AAB18463;  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX  
 DE A human TANGO 261 polypeptide clone.  
 XX  
 KW TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW edema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200052022-A1.  
 XX  
 PD 08-SEP-2000.  
 XX  
 PF 01-MAR-2000; 2000WO-US05226.  
 XX  
 PR 01-MAR-1999; 99US-0122458.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 XX  
 DR WPI; 2000-579269/54.  
 DR N-PSDB; AAA75165.  
 XX  
 PT Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX  
 PS Disclosure; Page -; 175pp; English.  
 XX  
 CC AAB18461-63 represent human TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or edema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema,  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammation,  
 CC bacterial and viral meningitis, Alzheimer's Disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC  
 CC note: the present sequence does not appear in the specification; it was

CC created using information provided.

XX Sequence 252 AA;

Query Match 99.8%; Score 1345; DB 21; Length 252;  
 Best Local Similarity 99.6%; Pred. No. 2.3e-111;  
 Matches 251; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRRDLSSDGLWPS 60

DB 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRRDLSSDGLWPS 60

QY 61 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQREHFRFQPTVPYLYQHEIDLPTTISLSDG 120

DB 61 ESTVSGNGIPEPQVYAPPRDRLAVPPFAQREHFRFQPTVPYLYQHEIDLPTTISLSDG 120

QY 121 EEPYPYQGFCTQLQRLDPEQQLNLNRESVRAPPNRTIFDSLDLMSARLGCPSPSSNSGIS 180

DB 121 EEPYPYQGFCTQLQRLDPEQQLNLNRESVRAPPNRTIFDSLDLMSARLGCPSPSSNSGIS 180

QY 181 ATCVSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAI 240

DB 181 ATCVSGGRMEGPPPTTYSEVIGHYVPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAI 240

QY 241 SKEKDQKQHPL 252

DB 241 SKEKDQKQHPL 252

RESULT 8

AAU78231

ID AAU78231 standard; Protein; 287 AA.

XX AC AAU78231;

DT 05-JUN-2002 (first entry)

DE Human Mechanically Induced Vascular Receptor 1, MIVR-1.

XX Human; MIVR-1; Mechanically Induced Vascular; receptor;  
 KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;  
 KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
 KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
 KW heart failure.

XX Homo sapiens.

XX WO200216416-A2.

PD 28-FEB-2002.

XX 21-AUG-2001; 2001WO-US26089.

XX 22-AUG-2000; 2000US-227159P.

XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.

XX (PFIZ ) PFIZER INC.

XX Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;

XX WPI; 2002-280912/32.

XX N-PSDB; ASK12137.

XX Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 PT Receptor-1 polypeptide, useful for treating cardiovascular diseases

XX Claim 14; Page 88-89; 105pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences AI761441.1, AI594390,  
 CC NM\_004338 and AQ177461. Also included are expression vectors, host

CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
CC with a candidate agent, where the molecule is a nucleic acid molecule  
CC comprising MIVR-1, IEX-1, VDUP-1, BTG-2 and TIS-11d or its  
CC expression product, determining if the anti-apoptotic activity is  
CC modulated and thereby identifying a modulator. The cardiac cell anti-  
CC apoptotic molecules and nucleic acids of the invention are useful for  
CC treating, diagnosing and monitoring progression of such diseases and  
CC disorders as characterised by increased apoptotic cell-death of vascular  
CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
CC stroke, arteriosclerosis and heart failure. The present sequence  
CC represents human MIVR-1.

XX Sequence 287 AA;  
SQ Query Match 99.4%; Score 1340; DB 23; Length 287;  
Best Local Similarity 99.2%; Pred. No. 7.6e-111;  
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAELFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
DB : |||||  
36 ITELEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPS 95  
QY 61 ESTVSGNGIPERQVYAPRPTDRLAVPPFAQRERFHRFQTPYLYQHEIDLPTTISLSDG 120  
DB |||||  
96 ESTVSGNGIPERQVYAPRPTDRLAVPPFAQRERFHRFQTPYLYQHEIDLPTTISLSDG 155  
QY 121 BEPPYQPCPTQLQRLDPEQQLNRESVRAPNRTIFDSDLMSARLGCPSPSSNSGIS 180  
DB |||||  
156 BEPPYQPCPTQLQRLDPEQQLNRESVRAPNRTIFDSDLMSARLGCPSPSSNSGIS 215  
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGLRLHHTHIAPLESAAIW 240  
DB |||||  
216 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGLRLHHTHIAPLESAAIW 275  
QY 241 SKEKDQKQGHPL 252  
DB |||||  
276 SKEKDQKQGHPL 287

RESULT 9  
ID AAM38712 standard; Protein; 237 AA.  
XX AC AAM38712;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 1857.  
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.  
XX PR 14-SEP-2000; 2000US-0662191.  
XX PR 19-OCT-2000; 2000US-0693036.  
XX PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Aeundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
DR N-PSDB; AAI57868.  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX Example 3; SEQ ID NO 1857; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC C.N.S disorders  
CC C.N.S disorders  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
SQ Sequence 237 AA;  
Query Match 95.1%; Score 1282; DB 22; Length 237;  
Best Local Similarity 100.0%; Pred. No. 8.5e-106; Indels 0; Gaps 0;  
Matches 237; Conservative 0; Mismatches 0;  
QY 16 MMVMVVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPSSTVSGNGIPEQVY 75  
DB |||||  
1 MMVMVVITCLLSHYKLSARSFISRHSGRRRDLSSGCLWPSSTVSGNGIPEQVY 60  
QY 76 APPRPTDLAVPPFAQRERFHRFQTPYLYQHEIDLPTTISLSDGEEPPPYQGCTQLQR 135  
DB |||||  
61 APPRPTDLAVPPFAQRERFHRFQTPYLYQHEIDLPTTISLSDGEEPPPYQGCTQLQR 120  
QY 136 DPEQQLNRESVRAPNRTIFDSDLMSARLGCPSPSSNSGISATCYGSGRMEGPP 195  
DB |||||  
121 DPEQQLNRESVRAPNRTIFDSDLMSARLGCPSPSSNSGISATCYGSGRMEGPP 180  
QY 196 TYSEVIGHYPGSSFQHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKQKQGHPL 252  
DB |||||  
181 TYSEVIGHYPGSSFQHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKQKQGHPL 237  
RESULT 10  
ID AAM38712 standard; Protein; 274 AA.  
XX AC AAM38712;  
XX DT 05-JUN-2002 (first entry)  
XX DE Mouse Mechanically Induced Vascular Receptor 1, MIVR-1.  
XX KW Mouse; MIVR-1; Mechanically Induced Vascular Receptor 1;  
KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;  
KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
KW heart failure.  
XX OS Mus musculus.  
XX PN



Db 121 LUGGFCPPSSNGISATCYGSGRMGGPPPTTSEVIGHYPGSSFOHQSSGPPSLLEGTR 180  
QY 226 LHHTHIAPLESAAIWSKEDKQKHPL 252  
Db 181 LHHTHIAPLESAAIWSKEDKQKHPL 207  
RESULT 12  
ID AAB18450 standard; Protein; 217 AA.  
AC AAB18450;  
XX  
XX 15-JAN-2001 (first entry)  
XX  
XX A murine TANGO 261 polypeptide.  
XX  
XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.  
XX  
XX Mus sp.  
OS  
XX  
XX WO200052022-A1.  
PN  
XX  
XX 08-SEP-2000.  
PD  
XX  
XX 01-MAR-2000; 2000WO-US05226.  
PP  
XX  
XX 01-MAR-1999; 99US-0129458.  
PR  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
PA  
XX  
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
PI  
XX  
XX WPI; 2000-579269/54.  
DR  
XX  
XX N-PSDB; AAA75152.  
DR  
XX  
XX Novel human and murine secreted proteins designated TANGO 216, 261,  
PT 262, 266 and 267 useful as modulating agents of cellular processes,  
PT e.g. for treating cancer -  
XX  
XX Claim 8; Fig 6A-B; 175pp; English.  
PS  
XX  
XX The present sequence represents a murine TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate  
CC cellular adhesion. The proteins can be used to treat any von Willebrand  
CC factor-associated disorder, regulate extracellular matrix structuring,  
CC cellular adhesion, and cell trafficking and/or migration, modulate  
CC cellular interactions, modulate cell adhesion in proliferative  
CC disorders, such as cancer, modulate the proliferation, differentiation,  
CC and/or function of cells that appear in the bone marrow, and leukocytes,  
CC treat bone marrow, blood and hematopoietic associated diseases and  
CC disorders, atelectasis, pulmonary congestion or oedema, emphysema,  
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
CC disorders, spleen associated diseases, modulate renal disorders, treat  
CC cardiovascular disorders such as ischemic heart disease, modulate the  
CC proliferation, differentiation, and/or function of bone and cartilage  
CC cells and to treat bone and/or cartilage associated diseases or  
CC disorder. They may also be used to treat disorders associated with the  
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic  
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's  
CC Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,

CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
CC disorders.  
XX  
SQ Sequence 217 AA;  
Query Match 79.1%; Score 1066.5; DB 21; Length 217;  
Best Local Similarity 88.9%; Pred. NO. 1.1e-86;  
Matches 201; Conservative 7; Mismatches 9; Indels 9; Gaps 3;  
QY 27 LSHVXLSARSFISRHSGRRREDALSSGCLWPSESTVSGMGIPPEQVYAPRPTDRLAV 86  
Db 1 LSHVXLSARSFISRHSGRRREDALSSGCLWPSESTVSG-GMPEQVYAPRPTDRLAV 59  
QY 87 PPFQQRERFHRFPQTPYLYQHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQLNRE 146  
Db 60 PPFQQRERFHRFPQTPYLYQHEIDLPTTISLSDGEEPPPYQGPCTQLQRDPQQLNRE 116  
QY 147 SVRAPPNRTTIEDSLMDSARLGGPCPPSSNSGISATCYGSGRMGGPPPTTSEVIGHYPG 206  
Db 117 SVRAPPNRTTIEDSLMDSARLGGPCPPSSNSGISATCYGSGRMGGPPPTTSEVIGHYPG 176  
QY 207 SSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEDKQKHPL 252  
Db 177 SSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEDKQKHPL 217  
RESULT 13  
ID ABG77052 standard; Protein; 241 AA.  
XX  
XX AC ABG77052;  
XX  
XX 05-NOV-2002 (first entry)  
DT  
XX  
XX Prostate specific protein #22.  
DE  
XX  
XX Human; prostate specific nucleic acid; PSNA; prostate cancer;  
KW cytostatic; non-cancerous prostate disease; PSP; vaccine;  
KW prostate specific protein; metastasis.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2000242776-A2.  
FN  
XX  
XX 30-MAY-2002.  
PD  
XX  
XX 01-NOV-2001; 2001WO-US45654.  
PP  
XX  
XX 01-NOV-2000; 2000US-244782P.  
PR  
XX  
XX (DIAD-) DIADEXUS INC.  
PA  
XX  
XX Sun Y, Recipon H, Chen S, Liu C;  
PI  
XX  
XX WPI; 2002-490217/52.  
DR  
XX  
XX New polypeptide useful for diagnosing and monitoring the presence and  
PT metastases of prostate cancer in a patient and as a component in  
PT databases for search analysis as well as in sequence analysis  
PT algorithms -  
XX  
XX Claim11; Page 213-214; 242pp; English.  
PS  
XX  
XX The invention relates to an isolated polypeptide comprising a sequence  
CC with 60 % identity to one of prostate specific protein (PSP) sequences,  
CC or comprising an amino acid sequence encoded by one of 136 nucleotide  
CC prostate specific nucleic acid sequences, PSNA, for a sequence that  
CC hybridises to it or is 60% identical to it), given in the specification.  
CC Also included are a vector comprising the polynucleotide, a host cell  
CC comprising the vector, an antibody specific for the PSP proteins and a  
CC vaccine comprising the protein or polynucleotide. The PSP and PSNA are  
CC useful for diagnosing and monitoring the presence and metastases of  
CC prostate cancer in a patient. The PSNA is useful for determining the

CC the level PSNA in a sample. An antibody to the PSP is useful for  
 CC determining the presence of prostate specific protein in a sample, and  
 CC for treating a patient with prostate cancer, which induces an immune  
 CC response against the prostate cancer cell expressing the nucleic acid or  
 CC polypeptide and a kit is useful for detecting a risk of cancer or  
 CC presence of cancer in a patient. PSNA is useful as hybridisation probes  
 CC to detect, characterise and quantify hybridising nucleic acids from both  
 CC genomic and transcript-derived nucleic acid samples and also in  
 CC microarrays. Sequences of PSP and PSNA are useful as components in  
 CC databases for search analysis as well as in sequence analysis algorithms.  
 CC PSNA is useful to drive in vivo expression of PSP. The present  
 CC sequence is a PSP of the invention.

XX Sequence 241 AA;

Query Match 79.0%; Score 1065.5; DB 23; Length 241;  
 Best Local Similarity 82.9%; Pred. No. 1.5e-86;  
 Matches 199; Conservative 0; Mismatches 0; Indels 41; Gaps 1;  
 QY 54 ECLNPSTVSGNGIPE-----P 72  
 DB 1 ECLNPSTVSGNGIPECCWPCDPPCRSSAPGSPALCSUHTGARTLPLFGGRP 60  
 QY 73 QYVAPRPDRLAVPPFAQRERFHFQPTYPYLOHEIDLPTTISLSDGEEPPPYQGPCTL 132  
 DB 61 QYVAPRPDRLAVPPFAQRERFHFQPTYPYLOHEIDLPTTISLSDGEEPPPYQGPCTL 120  
 QY 133 QLRDPEQLELNRESVRAPPNTTIFDSDLMDSARLGCGPPSSNGISATCYGSGRMEG 192  
 DB 121 QLRDPEQLELNRESVRAPPNTTIFDSDLMDSARLGCGPPSSNGISATCYGSGRMEG 180  
 QY 193 PPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPIESAAIWSKEKOKGHPL 252  
 DB 181 PPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPIESAAIWSKEKOKGHPL 240

RESULT 14

AAB18464  
 ID AAB18464 standard; Protein; 217 AA.

XX AAB18464;

XX 15-JAN-2001 (first entry)

DE A murine TANGO 261 polypeptide clone.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

XX Mus sp.

XX WO200052022-A1.

XX 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

XX 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Frazer CC;

XX WPI; 2000-579269/54.

DR N-PSDB; AAA75166.

XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 PT 262, 266 and 267 useful as modulating agents of cellular processes,  
 PT e.g. for treating cancer -  
 XX Disclosure; Page -; 175pp; English.

XX AAB18464-66 represent murine TANGO 261 proteins. The specification also  
 CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 CC polypeptides can be used to modulate cellular proliferation, modulate  
 CC cellular differentiation and/or modulate cellular adhesion. The  
 CC proteins can be used to treat any von Willebrand factor-associated  
 CC disorder, regulate extracellular matrix structuring, cellular adhesion,  
 CC and cell trafficking and/or migration, modulate cellular interactions,  
 CC modulate cell adhesion in proliferative disorders, such as cancer,  
 CC modulate the proliferation, differentiation, and/or function of cells  
 CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 CC and hematopoietic associated diseases and disorders, atelectasis,  
 CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 CC asthma and bronchiectasis, intestinal disorders, spleen associated  
 CC diseases, modulate renal disorders, treat cardiovascular disorders such  
 CC as ischemic heart disease, modulate the proliferation, differentiation,  
 CC and/or function of bone and cartilage cells and to treat bone and/or  
 CC cartilage associated diseases or disorder. They may also be used to  
 CC treat disorders associated with the ovaries, and cerebral oedema.  
 CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 CC bacterial and viral meningitis, Alzheimer's Disease, cerebral  
 CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 CC hydrocephalus and encephalitis, and treat hepatic disorders.  
 CC note: the present sequence does not appear in the specification; it was  
 CC created using information provided.

XX Sequence 217 AA;

Query Match 78.9%; Score 1063.5; DB 21; Length 217;  
 Best Local Similarity 88.5%; Pred. No. 1.9e-86;  
 Matches 200; Conservative 8; Mismatches 9; Indels 9; Gaps 3;

QY 27 LSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSGNGIPQVYAPRPDRLAV 86  
 DB 1 LSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSG-GMPEPQVYAPRPDRLAV 59  
 QY 87 PPEAQRERFHFQPTYPYLOHEIDLPTTISLSDGEEPPPYQGPCTLQLRDPEQLELNRE 146  
 DB 60 PPFQIRS---RFQPTYPYLOHEIALPTTISLSDGEEPPPYQGPCTLQLRDPEQLELNRE 116  
 QY 147 SVRAPPNRTIFDSDLMDSARLGCGPPSSNGISATCYGSGRMEGPPPTYSEVIGHYPG 206  
 DB 117 SVRAPPNRTIFDSDLDSTMLGCGPPSSNGISATCYGSGRMEGPPPTYSEVIGHYPG 176  
 QY 207 SSFOHQSSGPPSLLEGTRLHHTHIAPIESAAIWSKEKOKGHPL 252  
 DB 177 SSFOHQSSGPPSLLEGTRLHHTHIAPIESAAIWSKEKOKGHPL 217

RESULT 15

AAB18465  
 ID AAB18465 standard; Protein; 217 AA.

XX AAB18465;

XX 15-JAN-2001 (first entry)

DE A murine TANGO 261 polypeptide clone.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;



**THIS PAGE BLANK (U8PT0)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 03:11:01 ; Search time 4129 Seconds  
(without alignments)  
10512.258 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 2888711 seqs, 2045481396 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hg\_hum.\*  
31: em\_hg\_inv.\*  
32: em\_hg\_other.\*  
33: em\_hg\_mus.\*  
34: em\_hg\_pln.\*  
35: em\_hg\_rtd.\*  
36: em\_hg\_mam.\*  
37: em\_hg\_vrt.\*  
38: em\_sy.\*  
39: em\_hggo\_hum.\*  
40: em\_hggo\_mus.\*  
41: em\_hggo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	953	89.8	1061	9	BC015918	Homo sapi
2	926	87.3	1141	9	AF224278	Homo sapi
3	903	85.1	1818	9	AY128643	Homo sapi
4	902	85.0	4839	9	AF305616	Homo sapi
5	749	70.6	1321	6	AX392417	Sequence
6	701	66.1	861	6	AX392419	Sequence
7	695	65.5	61505	9	AF305426	Homo sapi
8	695	65.5	130435	9	HS718J7	Human DNA
9	509	48.0	1583	6	AX593655	Sequence
10	350	33.0	408	6	AX071267	Sequence
11	322	30.3	812	6	AX011709	Sequence
12	313	29.5	693	6	AX392430	Sequence
13	157	14.8	150224	9	HSJ1059L7	Human DNA
14	79	7.4	2570	9	AK056098	Homo sapi
15	51	4.8	51	6	AX199565	Sequence
16	44	4.1	651	10	AF220208	Mus muscu
17	44	4.1	878	6	AX392428	Sequence
18	44	4.1	1379	10	BC036995	Mus muscu
19	44	4.1	156698	10	AL837509	Mouse DNA
20	44	4.1	175754	2	AC110189	Mus muscu
21	44	4.1	176458	2	AC134911	Mus muscu
22	44	4.1	176821	2	AL837520	Mus muscu
23	42	4.0	249554	2	AX525744	Sequence
24	42	4.0	258632	2	AC139417	Rattus no
25	42	4.0	258632	2	AC111878	Rattus no
26	32	3.0	756	10	AB045588S2	Rattus no
27	32	3.0	56475	9	AL353658	Human DNA
28	32	3.0	110000	2	AC109084_1	Continuation (2 of
29	32	3.0	161462	2	AC136669	Rattus no
30	32	3.0	172525	9	AC008514	Homo sapi
31	32	3.0	179372	2	AC040922	Homo sapi
32	32	3.0	190025	2	AC140145	Homo sapi
33	32	3.0	190149	2	AC122097	Rattus no
34	32	3.0	192183	2	AC134492	Rattus no
35	32	3.0	198582	9	AC005291	Homo sapi
36	32	3.0	220084	2	AC123651	Mus muscu
37	32	3.0	221080	2	AC106160	Rattus no
38	32	3.0	226791	2	AC094199	Rattus no
39	32	3.0	239018	2	AC106202	Rattus no
40	32	3.0	242543	2	AC096128	Rattus no
41	32	3.0	249703	2	AC126988	Rattus no
42	32	3.0	258545	2	AC135666	Rattus no
43	32	3.0	259335	2	AC126194	Rattus no
44	32	3.0	279534	2	AC114035	Rattus no
45	32	3.0	281747	2	AC094865	Rattus no

ALIGNMENTS

RESULT 1  
BC015918  
LOCUS BC015918 1061 bp mRNA linear PRI 17-OCT-2001  
DEFINITION Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.  
ACCESSION BC015918  
VERSION BC015918.1 GI:16198474  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1061)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian











```
QY 372 ACCGGAGGGCTTCCACCGCTTCCAGCCCTTCCAGCCACCTATCCGTACCTGCGAGCACGAGATCGACC 431
|||
Db 57363 ACGGGAGCGCTTCCACCGCTTCCAGCCCTTCCAGCCACCTATCCGTACCTGCGAGCACGAGATCGACC 57422
|||
QY 432 TCGCGCCACCATCTCGCTGTGTCAGACGGGAGAGCCGCCACCTACCCCTACAGGGCCCTCGCA 491
|||
Db 57423 TCCACCCACCATCTCGCTGTGTCAGACGGGAGAGCCGCCACCTACCCCTACAGGGCCCTCGCA 57482
|||
QY 492 CCTCCAGCTTCGGGACCCCGGAGCAGCTGGAAGTGAACCGGAGTGGTGGCGGCAC 551
|||
Db 57483 CCTCCAGCTTCGGGACCCCGGAGCAGCTGGAAGTGAACCGGAGTGGTGGCGGCAC 57542
|||
QY 552 CCCCCAACAGAACCATCTTCGACAGTGACCTGATGATAGTGCCTAGGCTGGCGGCCCT 611
|||
Db 57543 CCCCCAACAGAACCATCTTCGACAGTGACCTGATGATAGTGCCTAGGCTGGCGGCCCT 57602
|||
QY 612 GCCCCCCCAGCAGTAGTAACCTCGGCGCATCAGCGCCACGTGCTACGGCAGCGCGGCGCATGG 671
|||
Db 57603 GCCCCCCCAGCAGTAGTAACCTCGGCGCATCAGCGCCACGTGCTACGGCAGCGCGGCGCATGG 57662
|||
QY 672 AGGGCGCGCGCCCGCCCTACAGCAGGTGATCGGCCACTACCCGGGGTCTCTTCCAGC 731
|||
Db 57663 AGGGCGCGCGCCCGCCCTACAGCAGGTGATCGGCCACTACCCGGGGTCTCTTCCAGC 57722
|||
QY 732 ACCACAGCAGTAGTGGCGCCCTCTTGTCTGGAGGGACCCGGCTCCACACACACACA 791
|||
Db 57723 ACCACAGCAGTAGTGGCGCCCTCTTGTCTGGAGGGACCCGGCTCCACACACACACA 57782
|||
QY 792 TCGCGCCCTACAGAGCGCAGCCATCTGGGACCAAGAGAGATTAACAGAAAGCACACC 851
|||
Db 57783 TCGCGCCCTACAGAGCGCAGCCATCTGGGACCAAGAGAGATTAACAGAAAGCACACC 57842
|||
QY 852 CTCTCTAGGGTCCCGAGGGGCGCGGGTGGGGCTGCGTAGTGTAAGGAGGAGACATC 911
|||
Db 57843 CTCTCTAGGGTCCCGAGGGGCGCGGGTGGGGCTGCGTAGTGTAAGGAGGAGACATC 57902
|||
QY 912 CCGCTTCTTAGAAGAGGTGAGAGGAAGCGGGGGCGCGAGCAACCGATCGTGGCC 971
|||
Db 57903 CCGCTTCTTAGAAGAGGTGAGAGGAAGCGGGGGCGCGAGCAACCGATCGTGGCC 57962
|||
QY 972 CTCCTCCCTCCACCTCCCTGTGTATTAATTTACATGTGATCTGCTGTGATGCACAA 1031
|||
Db 57963 CTCCTCCCTCCACCTCCCTGTGTATTAATTTACATGTGATCTGCTGTGATGCACAA 58022
|||
QY 1032 GCTAAGAGAGCTTGCAAAAAA 1057
|||
Db 58023 GCTAAGAGAGCTTGCAAAAAA 58048
|||
```

```
RESULT 8
HS718J7/c
LOCUS
DEFINITION
Human DNA sequence from clone RP4-718J7 on chromosome
20q13.31-13.33 Contains the PKC1 gene for soluble
phosphoenolpyruvate carboxykinase 1, part of a novel gene similar
to mouse DLM-1 (tumour stroma and activated macrophage protein),
the 3' end of the TMEPAI gene encoding an androgen induced lb
transmembrane protein (PMEPAI), two putative novel genes, a Cpg
island, ESTs, STSs and GSSs, complete sequence.
ACCESSION
AL035541
VERSION
AL035541.15 GI:11546043
KEYWORDS
HTG; Cpg Island; DLM-1; macrophage protein; PKC1;
phosphoenolpyruvate carboxykinase; PMEPAI; TMEPAI; transmembrane
protein.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 130435)
AUTHORS
Sehara,H.
TITLE
Direct Submission
JOURNAL
Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
```

## COMMENT

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Dec 5, 2000 this sequence version replaced gi:10196628.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr20  
This sequence is the entire insert of clone RP4-718J7 The true left end of clone RP3-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579F20 is at 43945 in this sequence.  
RP4-718J7 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCRPAC2

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## FEATURES

```
source
1..130435
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.31-13.33"
/clone="RP4-718J7"
/clone_lib="RPCI-4"
repeat_region 319..478
/note="5 copies 32 mer 67% conserved"
repeat_region 370..463
/note="2 copies 47 mer 87% conserved"
repeat_region 427..626
/note="4 copies 50 mer 71% conserved"
repeat_region 707..812
/note="53 copies 2 mer cc 61% conserved"
repeat_region 967..2752
/note="893 copies 2 mer gg 54% conserved"
repeat_region 982..2730
/note="33 copies 53 mer 54% conserved"
repeat_region 1177..1820
/note="4 copies 161 mer 64% conserved"
repeat_region 1204..1371
/note="3 copies 56 mer 75% conserved"
repeat_region 1358..2757
/note="28 copies 50 mer 54% conserved"
repeat_region 1675..2718
/note="18 copies 58 mer 55% conserved"
repeat_region 1843..2346
/note="9 copies 56 mer 64% conserved"
repeat_region 1852..2079
/note="4 copies 57 mer 86% conserved"
repeat_region 2028..2632
/note="11 copies 55 mer 60% conserved"
repeat_region 2087..2257
/note="3 copies 57 mer 83% conserved"
repeat_region 2228..2497
/note="5 copies 54 mer 75% conserved"
repeat_region 2583..2716
/note="2 copies 67 mer 82% conserved"
```

```
repeat_region 3378. .3426
/note="L1MB4 repeat: matches 6088. .6136 of consensus"
repeat_region 3799. .3896
/note="Charliel repeat: matches 681. .781 of consensus"
repeat_region 5331. .5793
/note="MLT1B repeat: matches 14. .466 of consensus"
repeat_region 5797. .5988
/note="6 copies 32 mer 86% conserved"
repeat_region 5855. .5978
/note="31 copies 4 mer gcac 61% conserved"
repeat_region 6474. .6591
/note="MIR repeat: matches 91. .218 of consensus"
repeat_region 6592. .6723
/note="FLAM_C repeat: matches 1. .132 of consensus"
repeat_region 6952. .7021
/note="L2 repeat: matches 2637. .2705 of consensus"
repeat_region 7358. .7671
/note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 8521. .8554
/note="Alu repeat: matches 1. .34 of consensus"
repeat_region 8863. .9217
/note="MLT1D repeat: matches 105. .505 of consensus"
repeat_region 9228. .9746
/note="L2 repeat: matches 1459. .2028 of consensus"
repeat_region 9880. .10089
/note="L2 repeat: matches 2292. .2511 of consensus"
repeat_region 10094. .10206
/note="AluDb repeat: matches 188. .300 of consensus"
repeat_region 10263. .10573
/note="AluJb repeat: matches 1. .308 of consensus"
repeat_region 10658. .10753
/note="L2 repeat: matches 2647. .2749 of consensus"
repeat_region 11816. .12380
/note="LTR19B repeat: matches 1. .580 of consensus"
repeat_region 12519. .12813
/note="L1MD2 repeat: matches 6032. .6331 of consensus"
repeat_region 13559. .13726
/note="3 copies 56 mer 73% conserved"
repeat_region 13585. .13716
/note="22 copies 6 mer ctctct 67% conserved"
repeat_region 13586. .13717
/note="66 copies 2 mer tc 68% conserved"
repeat_region 13588. .13715
/note="4 copies 32 mer 78% conserved"
repeat_region 13589. .13724
/note="34 copies 4 mer cttt 77% conserved"
misc_feature complement(13622. .14142)
/note="match: GSS: Em:AQ592603"
repeat_region 14313. .14437
/note="LTR16C repeat: matches 257. .386 of consensus"
repeat_region 15015. .15110
/note="3 copies 32 mer 79% conserved"
repeat_region 15021. .15110
/note="45 copies 2 mer ac 80% conserved"
repeat_region 15026. .15109
/note="7 copies 12 mer 82% conserved"
repeat_region 15028. .15111
/note="14 copies 6 mer cacaca 82% conserved"
repeat_region 15029. .15108
/note="20 copies 4 mer acac 83% conserved"
repeat_region 15273. .15399
/note="L2 repeat: matches 2159. .2285 of consensus"
repeat_region 15420. .15721
/note="AluSx repeat: matches 1. .303 of consensus"
repeat_region 16333. .16396
/note="2 copies 32 mer 98% conserved"
repeat_region 16575. .16690
/note="L2 repeat: matches 2596. .2711 of consensus"
repeat_region 18049. .18169
/note="MIR repeat: matches 86. .211 of consensus"
repeat_region 18312. .18438
/note="MIR repeat: matches 138. .250 of consensus"
repeat_region 18585. .18776
```

```
repeat_region 19211. .19290
/note="MER20 repeat: matches 7. .217 of consensus"
repeat_region 20413. .20463
/note="MIR repeat: matches 72. .154 of consensus"
misc_feature complement(21005. .21672)
/note="match: GSS: Em:AQ748384"
misc_feature complement(21187. .21672)
/note="match: GSS: Em:AQ776209"
misc_feature complement(21278. .21663)
/note="match: GSS: Em:AQ136459"
misc_feature complement(21392. .21672)
/note="match: GSS: Em:AQ892114"
repeat_region 21667. .21775
/note="MIR repeat: matches 73. .192 of consensus"
repeat_region 23577. .23646
/note="2 copies 35 mer 100% conserved"
repeat_region 24257. .24591
/note="L2 repeat: matches 2336. .2710 of consensus"
repeat_region 25580. .25611
/note="16 copies 2 mer tc 87% conserved"
repeat_region 26333. .26643
/note="AluY repeat: matches 1. .310 of consensus"
repeat_region 27603. .27684
/note="L2 repeat: matches 2661. .2739 of consensus"
repeat_region 27823. .28041
/note="MIR repeat: matches 38. .242 of consensus"
repeat_region 28129. .28259
/note="FLAM_C repeat: matches 1. .127 of consensus"
repeat_region 31203. .31356
/note="L1MB5 repeat: matches 6015. .6176 of consensus"
repeat_region 32224. .32522
/note="AluSx repeat: matches 1. .298 of consensus"

Query Match 65.5%; Score 695; DB 9; Length 130435;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 312 AGCGCAGGTCTACGCCCGCGCTCGGCCACCGACGGCTGGCCCTGGCCCTTCGCC 371
DB 128443 AGCGCAGGTCTACGCCCGCGCTCGGCCACCGACGGCTGGCCCTTCGCC 128384

QY 372 AGCGGGAGCGTTCACCGCTTCCAGCCACCTATCCGTACTCTGACGACGATCGACC 431
DB 128393 AGCGGGAGCGTTCACCGCTTCCAGCCACCTATCCGTACTCTGACGACGATCGACC 128324

QY 432 TGCGCCGCCACCATCTCGTGTGTCAGACGGGGAGGAGCCCCACCTTACCAGGCCCTGCA 491
DB 128323 TGCCACCCACCATCTCGTGTGTCAGACGGGGAGGAGCCCCACCTTACCAGGCCCTGCA 128264

QY 492 CCCTCCAGTTCGGGACCCCGGACGACGCTGGAACTGAACCGGAGTGGTTCGGGCAC 551
DB 128263 CCCTCCAGTTCGGGACCCCGGACGACGCTGGAACTGAACCGGAGTGGTTCGGGCAC 128204

QY 552 CCCCAACAGAACCATCTTCGACAGTGACCTGATGATAGTCCAGGCTGGCGGCCCT 611
DB 128203 CCCCAACAGAACCATCTTCGACAGTGACCTGATGATAGTCCAGGCTGGCGGCCCT 128144

QY 612 GCGCCCGCCAGAGTAATCTCGGCGATCAGCGCCACCTGCTACGGCAGCGCGCGCATGG 671
DB 128143 GCGCCCGCCAGAGTAATCTCGGCGATCAGCGCCACCTGCTACGGCAGCGCGCGCATGG 128084

QY 672 AGGGCCCGCCCGCCACCTTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTCTCCAGC 731
DB 128083 AGGGCCCGCCCGCCACCTTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTCTCCAGC 128024

QY 732 ACCAGCAGCAGTGGCGCGCTCTCTTGTGGAGGGAACCGGCTCCACACACACACA 791
DB 128023 ACCAGCAGCAGTGGCGCGCTCTCTTGTGGAGGGAACCGGCTCCACACACACACA 127964

QY 792 TCGCGCCCTTAGAGAGCGCAGCCCATCTCGAGCAAGAGAGATTAACAGAAAGACACC 851
DB 127963 TCGCGCCCTTAGAGAGCGCAGCCCATCTCGAGCAAGAGAGATTAACAGAAAGACACC 127904
```





RESULT 11  
AX011709  
LOCUS AX011709 812 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 107 from Patent WO9955858.  
ACCESSION AX011709  
VERSION AX011709.1 GI:9998233  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and  
Pilarczyk, C.  
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue  
JOURNAL Patent: WO 9955858-A 107 04-NOV-1999;  
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN  
(DE); PILARSKY CHRISTIAN (DE)  
FEATURES  
source Location/Qualifiers  
1..812  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 157 a 272 c 237 g 146 t  
ORIGIN  
Query Match 30.3%; Score 322; DB 6; Length 812;  
Best Local Similarity 100.0%; Pred. No. 7e-157;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 168 TCACGTGCTGTGAGCCACTACAGCTGTCTGCACGGTCTTTCATCAGCGGCACAGCC 227  
DB 426 TCACGTGCTGTGAGCCACTACAGCTGTCTGCACGGTCTTTCATCAGCGGCACAGCC 485  
QY 228 AGGGCGGAGGAGAGATGCCCTGCTCAGAGGATGCTGTGGCCCTCGGAGACA 287  
DB 486 AGGGCGGAGGAGAGATGCCCTGCTCAGAGGATGCTGTGGCCCTCGGAGACA 545  
QY 288 CAGTGTACGCAACCGAATCCAGAGCGCGAGTCTAGCGCCCGCTCGGCCACCGACC 347  
DB 546 CAGTGTACGCAACCGAATCCAGAGCGCGAGTCTAGCGCCCGCTCGGCCACCGACC 605  
QY 348 GCCTGGCGTGGCGCCCTTGGCCAGCGGAGCGCTTCACCGCTTCAGGCCACCTATC 407  
DB 606 GCCTGGCGTGGCGCCCTTGGCCAGCGGAGCGCTTCACCGCTTCAGGCCACCTATC 665  
QY 408 CGTACCTGCAGCAGAGATCGACCTGCGCCGCCACCATCTCGCTGTCAGACGGGAGGAGC 467  
DB 666 CGTACCTGCAGCAGAGATCGACCTGCGCCGCCACCATCTCGCTGTCAGACGGGAGGAGC 725  
QY 468 CCCACCCCTACCGAGGCCCTG 489  
DB 726 CCCACCCCTACCGAGGCCCTG 747  
RESULT 12  
AX392430/c  
LOCUS AX392430 693 bp DNA linear PAT 23-MAR-2002  
DEFINITION Sequence 14 from Patent WO0216416.  
ACCESSION AX392430  
VERSION AX392430.1 GI:19700746  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and  
Turi, T.G.  
TITLE Diagnosis and treatment of cardiovascular conditions

JOURNAL Patent: WO 0216416-A 14 28-FEB-2002;  
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)  
FEATURES  
source Location/Qualifiers  
1..693  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 101 a 205 c 237 g 149 t 1 others  
ORIGIN  
Query Match 29.5%; Score 313; DB 6; Length 693;  
Best Local Similarity 99.3%; Pred. No. 3.6e-152;  
Matches 583; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 468 CCCACCCCTACCGAGGCCCTGCGACCTTCAGCTTCGGGACCCCGAGCAGCTGGGAAC 527  
DB 587 CCCACCCCTACCGAGGCCCTGCGACCTTCAGCTTCGGGACCCCGAGCAGCTGGGAAC 528  
QY 528 TGAACCGGAGTCCGTGCGCGCACCCCAACAGAACCATCTTCACAGTGACCTGATGG 587  
DB 527 TGAACCGGAGTCCGTGCGCGCACCCCAACAGAACCATCTTCACAGTGACCTGATGG 468  
QY 588 ATAGTGCAGCTGGGCGGCCCTGCGCCGCCAGCAGTAACCTCGGGCATCAGCGCACGT 647  
DB 467 ATAGTGCAGCTGGGCGGCCCTGCGCCGCCAGCAGTAACCTCGGGCATCAGCGCACGT 408  
QY 648 GCTAGGCGAGCGGCGGCGCATGAGGGCGCGGCCACCTACAGCAGGTCTCGGCC 707  
DB 407 GCTAGGCGAGCGGCGGCGCATGAGGGCGCGGCCACCTACAGCAGGTCTCGGCC 348  
QY 708 ACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGCGCGCGCCCTCTTGTCTGGA-G 766  
DB 347 ACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGCGCGCGCCCTCTTGTCTGGA 288  
QY 767 GGGACCGCGCTCCACCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAA 826  
DB 287 GGGACCGCGTTCCCCCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAA 228  
QY 827 GAGAAGGATAACAGAAAGGACACCTCTTAGGGTCCCGAGGGGGCGGGCTGGGCT 886  
DB 227 GAGAAGGATAACAGAAAGGACACCTCTTAGGGTCCCGAGGGGGCGGGCTGGGCT 168  
QY 887 GCGTAGGTGAAAAGGAGAACACCTCCGGCTCTTTAGAAAGAGGAGTGAGAGAGCGGG 946  
DB 167 GCGTAGGTGAAAAGGAGAACACCTCCGGCTCTTTAGAAAGAGGAGTGAGAGAGCGGG 108  
QY 947 GGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTTATAAATTATTA 1006  
DB 107 GGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTTATAAATTATTA 48  
QY 1007 TGTGATGCTGCTGTAATGCACAGCTTAAGAGAGCTTGCAAAAAA 1053  
DB 47 TGTGATGCTGCTGTAATGCACAGCTTAAGAGAGCTTGCAAAAAA 1  
RESULT 13  
HSJ1059L7/c  
LOCUS HSJ1059L7 150224 bp DNA linear PRI 24-FEB-2001  
DEFINITION Human DNA sequence from clone RP5-1059L7 on chromosome  
20q13.2-13.33 Contains the 5' end of the TMEPAI (PMEPAI) gene  
encoding an androgen induced 1b transmembrane protein, ESTs, STSs,  
GSSs and two CpG islands, complete sequence.  
ACCESSION AL121913  
VERSION AL121913.4 GI:7161781  
KEYWORDS HTG; CpG island; PMEPAI; TMEPAI; transmembrane protein.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (Bases 1 to 150224)  
AUTHORS Skuce, C.  
TITLE Direct Submission  
JOURNAL Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone  
requests: clonerquest@sanger.ac.uk  
On Mar 6, 2000 this sequence version replaced gi:7007305.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 20, constructed by the Sanger Centre Chromosome 20  
Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>  
This sequence is the entire insert of clone RP5-1059L7. The true  
left end of clone RP11-402F1 is at 106677 in this sequence. This  
sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. RP5-1059L7 is from  
the library RPCI-5 constructed by the group of Pieter de Jong. For  
further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: PCVPAC2.  
Location/Qualifiers  
1..150224  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="q13.2-13.33"  
/clone="RP5-1059L7"  
/clone\_lib="RPCI-5"  
283..355  
/note="MIR repeat: matches 71..145 of consensus"  
repeat\_region  
/note="MIR repeat: matches 71..145 of consensus"  
repeat\_region  
1509..1628  
/note="MIR repeat: matches 46..165 of consensus"  
repeat\_region  
2076..2155  
/note="L2 repeat: matches 2616..2707 of consensus"  
gene  
complement(2445..53425)  
/gene="TMEPAI"  
mRNA  
complement(join(<2445..2599,53362..53425))  
/gene="TMEPAI"  
/product="dJ1059L7.1.2 (androgen induced type 1b  
transmembrane protein (PMEPA1), isoform 2)"  
/note="match: cDNAs: Em:AF224278  
match: ESTs: Em:AA088767"  
/evidence=not\_experimental  
complement(join(<2445..2599,53362..53365))  
/gene="TMEPAI"  
CDS  
/note="continues in dJ718J7 (AL035541)"  
/codon\_start=1  
/evidence=not\_experimental  
/product="dJ1059L7.1.1 (androgen induced type 1b  
transmembrane protein (PMEPA1), isoform 2)"  
/protein\_id="CAC32857.1"  
/db\_xref="GI:13160408"  
/translation="MAELEFVQIIIVVMVMVITCLLSHYKLSARSPISRHSQ  
RRREDALSS"  
complement(join(<2445..2599,52376..52477))  
/gene="TMEPAI"  
mRNA  
/product="dJ1059L7.1.1 (androgen induced type 1b  
transmembrane protein (PMEPA1), isoform 1)"  
/note="match: cDNAs: Em:AF009426 Em:AF009427 Em:AF009425  
Em:AF009424  
match: ESTs: Em:AA249792 Em:AI594390 Em:AA128075  
Em:AA088767"  
/evidence=not\_experimental  
complement(join(<2445..2599,52376..52472))  
/gene="TMEPAI"  
/note="continues in dJ718J7 (AL035541)"  
match: proteins: Tr:O15166 Tr:O15168"  
/codon\_start=1  
/evidence=not\_experimental  
/product="dJ1059L7.1.1 (androgen induced type 1b  
transmembrane protein (PMEPA1), isoform 1)"  
/protein\_id="CAB88144.1"  
/db\_xref="GI:7619746"  
/db\_xref="SPTREMBL:Q9NTR9"  
/translation="MGVNSTAAARAAQPNVSCNCKRSLFQSMTELEFVQIIIV  
VVMVMVITCLLSHYKLSARSPISRHSQRRREDALSS"  
2538..2567  
/note="10 copies 3 mer cac 90% conserved"  
3128..3606  
/note="match: GSS: Em:AQ059507"  
3214..3324  
/note="3 copies 37 mer 83% conserved"  
3339..3413  
/note="3 copies 25 mer 92% conserved"  
3390..3921  
/note="19 copies 28 mer 61% conserved"  
3394..3893  
/note="25 copies 20 mer 59% conserved"  
3407..3892  
/note="27 copies 18 mer 58% conserved"  
3421..3910  
/note="7 copies 70 mer 62% conserved"  
3473..3904  
/note="12 copies 36 mer 58% conserved"  
3553..3932  
/note="10 copies 38 mer 64% conserved"  
3573..3697  
/note="5 copies 25 mer 76% conserved"  
3746..3921  
/note="8 copies 22 mer 64% conserved"  
4032..4069  
/note="19 copies 2 mer tt 81% conserved"  
4792..5001  
/note="3 copies 70 mer 82% conserved"  
5002..5190  
/note="MER20 repeat: matches 1..200 of consensus"  
5224..5282  
/note="MIR repeat: matches 94..154 of consensus"  
8103..8205  
/note="MIR repeat: matches 48..161 of consensus"  
complement(8858..9277)  
/gene="TMEPAI"  
/note="match: GSS: Em:AQ881699"  
9224..9382  
/note="MIR repeat: matches 81..239 of consensus"  
9286..9706  
/note="match: GSS: Em:AQ565377"  
9295..9861  
/note="match: GSS: Em:AQ757992"  
10035..10140  
/note="Charlie4 repeat: matches 39..146 of consensus"  
11175..11343  
/note="MIR repeat: matches 11..196 of consensus"  
11630..12037  
/note="L2 repeat: matches 2290..2705 of consensus"  
12054..12362  
/note="Alusx repeat: matches 1..311 of consensus"  
14448..14493  
/note="23 copies 2 mer at 76% conserved"  
14634..14813  
/note="MER20 repeat: matches 1..187 of consensus"  
15096..15497  
/note="CpG island"  
/evidence=not\_experimental

RESULT 14  
AK056098/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AK056098 2570 bp mRNA linear  
Homo sapiens CDNA FLJ31536 fis, clone NT2R12000699.  
AK056098  
AK056098.1 GI:16551411  
oligo capping; fis (full insert sequence).  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1  
Ninomiya, K., Wazatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,  
Kodaïra, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, K., Abe, K.,  
Kanihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,  
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,

BASE COUNT 12 a 16 c 18 g 5 t

ORIGIN

Query Match 4.8%; Score 51; DB 6; Length 51;  
Best Local Similarity 100.0%; Pred. No. 7.4e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 CACGGTCCTTCATCAGCCGGGCACAGCCAGGGGGCGGAGAGAGATGCCC 251  
|||  
Db 1 CACGGTCCTTCATCAGCCGGGCACAGCCAGGGGGCGGAGAGAGATGCCC 51  
|||

Search completed: December 8, 2003, 05:33:47  
Job time : 4149 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 02:02:00 ; Search time 344 Seconds  
(without alignments)  
8325.883 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_19Jun03.*				
1:	/SIDS1/gcgdata/geneseq/geneq-emb1/NA1980.DAT.*			
2:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1981.DAT.*			
3:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1982.DAT.*			
4:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1983.DAT.*			
5:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1984.DAT.*			
6:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1985.DAT.*			
7:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1986.DAT.*			
8:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1987.DAT.*			
9:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1988.DAT.*			
10:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1989.DAT.*			
11:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1990.DAT.*			
12:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1991.DAT.*			
13:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1992.DAT.*			
14:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1993.DAT.*			
15:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1994.DAT.*			
16:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1995.DAT.*			
17:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1996.DAT.*			
18:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1997.DAT.*			
19:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1998.DAT.*			
20:	/SIDS1/gcgdata/geneq/geneq-emb1/NA1999.DAT.*			
21:	/SIDS1/gcgdata/geneq/geneq-emb1/NA2000.DAT.*			
22:	/SIDS1/gcgdata/geneq/geneq-emb1/NA2001A.DAT.*			
23:	/SIDS1/gcgdata/geneq/geneq-emb1/NA2001B.DAT.*			
24:	/SIDS1/gcgdata/geneq/geneq-emb1/NA2002.DAT.*			
25:	/SIDS1/gcgdata/geneq/geneq-emb1/NA2003.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1061	100.0	1061	21 AAA47429	Sequence encoding
2	1056	99.5	1334	25 ABZ35103	Human secretory po
3	961	90.6	969	21 AAA75151	cDNA encoding a hu
4	952	89.7	969	21 AAA75163	cDNA clone encodin
5	926	87.3	1140	24 ABK92120	Prostate cancer-as
6	910	85.8	969	21 AAA75164	cDNA clone encodin
7	910	85.8	969	21 AAA75165	cDNA clone encodin
8	891	84.0	1066	22 AAA157868	Human polynucleoti

9	749	70.6	1321	24 ABK12137	Human cDNA encodin
c 10	548	51.6	1069	22 AAI59654	Human polynucleoti
c 11	509	48.0	1583	24 ABS61424	Prostate specific
c 12	415	39.1	474	25 ABZ84732	Toxicologically re
c 13	350	33.0	408	22 AAF65983	Novel human polynu
c 14	322	30.3	812	20 AAZ52964	Human prostate tum
c 15	313	29.5	693	24 ABK12143	Human MIVR-1 homol
c 16	229	21.6	254	21 AAA41265	Human secreted exp
c 17	93	8.8	426	23 AAS84502	DNA encoding novel
c 18	91	8.6	1879	23 AAS84503	DNA encoding novel
c 19	68	6.4	522	24 ABT10027	Human breast cance
c 20	60	5.7	60	24 ABN40872	Human spliced tran
c 21	51	4.8	51	22 AAH89714	Human coding sequ
c 22	44	4.1	878	24 ABK12142	Mouse cDNA encodin
c 23	44	4.1	1713	21 AAA75152	cDNA encoding a mu
c 24	44	4.1	1713	21 AAA75166	cDNA clone encodin
c 25	44	4.1	1713	21 AAA75167	cDNA clone encodin
c 26	44	4.1	1713	21 AAA75168	cDNA clone encodin
c 27	42	4.0	673	24 ABT09178	Phase-1 Rat CT gen
c 28	32	3.0	577	22 ABA50270	Human breast cell
c 29	32	3.0	577	22 ABA68210	Human foetal liver
c 30	32	3.0	577	22 ABA35222	Probe #13688 for g
c 31	32	3.0	577	22 AAK16592	Human brain expres
c 32	32	3.0	577	22 AAK42346	Human bone marrow
c 33	32	3.0	577	22 AAI23118	Probe #13051 for g
c 34	32	3.0	577	22 AAI48425	Probe #17111 used
c 35	32	3.0	577	22 AAI08772	Probe #8763 used t
c 36	32	3.0	577	22 ABA41955	Human liver single
c 37	32	3.0	577	23 ABS41642	Human genome-deriv
c 38	32	3.0	1964	22 ABA45134	Human breast cell
c 39	32	3.0	1964	22 ABA55612	Human foetal liver
c 40	32	3.0	1964	22 ABA25310	Probe #3776 for ge
c 41	32	3.0	1964	22 AAK03841	Human brain expres
c 42	32	3.0	1964	22 AAK29307	Human bone marrow
c 43	32	3.0	1964	22 AAI13901	Probe #3834 for ge
c 44	32	3.0	1964	22 AAI35267	Probe #3953 used t
c 45	32	3.0	1964	22 AAI03769	Probe #3760 used t

ALIGNMENTS

RESULT 1  
AAAA7429  
ID AAA47429 standard; DNA; 1061 Bp.

AC AAA47429;

XX 20-OCT-2000 (first entry)

Sequence encoding human neuron-associated protein.

Neuron associated protein; NEUAP; neurological disorder; epilepsy; ischemic cerebrovascular disease; stroke; cerebral neoplasm; Alzheimer's disease; Pick's disease; Huntington's disease; dementia; Parkinson's disease; demyelinating disease; meningitis; prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy; muscular dystrophy; central nervous system; CNS; peripheral nervous system; PNS; myopathy; schizophrenia; actinic keratosis; arteriosclerosis; atherosclerosis; bursitis; cirrhosis; hepatitis; mixed connective tissue disease; MCTD; myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer; autoimmune disease; inflammation; acquired immunodeficiency syndrome; AIDS; Addison's disease; adult respiratory distress syndrome; allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma; Werner syndrome; trauma; human; ds.  
Homo sapiens.  
Location/Qualifiers  
101..859  
/\*tag= a  
/product= Neuron associated protein

XX	PN	WO200034477-A2.
XX	PD	15-JUN-2000.
XX	PF	10-DEC-1999; 99WO-US30408.
XX	PR	11-DEC-1998; 98US-0210083.
XX	PR	11-DEC-1998; 98US-9123456.
XX	PR	09-FEB-1999; 99US-0119365.
XX	PR	16-MAR-1999; 99US-0124687.
XX	PA	(INCY-) INCYTE PHARM INC.
XX	PI	Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;
XX	PI	Lu DAM, Azimzai Y;
XX	DR	WPI; 2000-423423/36.
XX	DR	P-PSDB; AAB01388.
XX	PT	New human neuron-associated proteins and polynucleotides encoding them,
XX	PT	useful for diagnosis, treatment and prevention of cell proliferative
XX	PT	disorders including cancer, neuronal and neurological disorders
XX	PS	Claim 9; Page 136; 145pp; English.
XX	CC	Human neuron-associated proteins (NEUP) can be used for for
XX	CC	treating or preventing a disorder associated with decreased
XX	CC	expression or activity of NEUP. Antagonists of NEUP are useful for
XX	CC	treating or preventing disorder associated with increased expression
XX	CC	or activity of NEUP. NEUP or their fragments or derivatives are
XX	CC	useful for treating neurological disorder such as epilepsy, ischemic
XX	CC	cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's
XX	CC	disease, Pick's disease, Huntington's disease, dementia and
XX	CC	Parkinson's disease. NEUPs are also useful for treating other
XX	CC	demyelinating diseases, bacterial and viral meningitis, prion
XX	CC	diseases including kuru, Creutzfeldt-Jakob disease, nutritional and
XX	CC	metabolic diseases of the nervous system, neurofibromatosis, other
XX	CC	developmental disorders of the central nervous system, cerebral
XX	CC	palsy, neuroskeletal disorders, autonomic nervous system disorders,
XX	CC	cranial nerve disorders, spinal cord diseases, muscular dystrophy and
XX	CC	other neuromuscular disorders, peripheral nervous system disorders,
XX	CC	inherited, metabolic, endocrine, and toxic myopathies, mental
XX	CC	disorders including mood, anxiety and schizophrenic disorders, a cell
XX	CC	proliferative disorder such as actinic keratosis, arteriosclerosis,
XX	CC	atherosclerosis, psoriasis, cirrhosis, hepatitis, mixed connective
XX	CC	tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal
XX	CC	haemoglobinuria, cancers of the adrenal gland, bladder, bone,
XX	CC	bone marrow, brain, breast, cervix, and an autoimmune/inflammatory
XX	CC	disorder such as acquired immunodeficiency syndrome (AIDS), Addison's
XX	CC	disease, adult respiratory distress syndrome, allergies, ankylosing
XX	CC	spondylitis, amyloidosis, anemia, asthma, Werner syndrome,
XX	CC	complications of cancer, hemodialysis, and extracorporeal circulation,
XX	CC	viral, bacterial, fungal parasitic, protozoal, and helminthic
XX	CC	infections, and trauma. This sequence was given the Incyte ID no.
XX	CC	1871288CB1.
XX	CC	Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;
XX	CC	Query Match 100.0%; Score 1061; DB 21; Length 1061;
XX	CC	Best Local Similarity 100.0%; Pred. No. 0;
XX	CC	Matches 1061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	TCCTCTTGGGTTGGGTTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
DB	1	TCCTCTTGGGTTGGGTTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
QY	61	TGAGAACTGAAGCGGACGGCTTCTCTCGAAACAGGCAATGGCGGAGCTGGAGTTGT 120
DB	61	TGAGAACTGAAGCGGACGGCTTCTCTCGAAACAGGCAATGGCGGAGCTGGAGTTGT 120
QY	121	TCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
DB	121	TCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180





```
XX AC AAA75151;
XX DT 15-JAN-2001 (first entry)
XX DE cDNA encoding a human TANGO 261 polypeptide.
XX KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 6..764
FT /*tag= a
FT /product= "TANGO 261"
FT sig_peptide 6..89
FT /*tag= b
FT mat_peptide 90..764
FT /*tag= c
XX WO200052022-A1.
XX PD 08-SEP-2000.
XX PF 01-MAR-2000; 2000WO-US05226.
XX PR 01-MAR-1999; 99US-0122458.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX WPI; 2000-579269/54.
XX P-PSDB; AAB18449.
XX Novel human and murine secreted proteins designated TANGO 216, 261,
XX 262, 266 and 267 useful as modulating agents of cellular processes,
XX e.g. for treating cancer -
XX Claim 2; Fig 5; 175pp; English.
XX The present sequence encodes a human TANGO 261 polypeptide. The
XX specification also describes TANGO 266, TANGO 216, TANGO 262, and
XX TANGO 267. The TANGO polypeptides can be used to modulate cellular
XX proliferation, modulate cellular differentiation and/or modulate
XX cellular adhesion. The proteins can be used to treat any von Willebrand
XX factor-associated disorder, regulate extracellular matrix structuring,
XX cellular adhesion, and cell trafficking and/or migration, modulate
XX cellular interactions, modulate cell adhesion in proliferative
XX disorders, such as cancer, modulate the proliferation, differentiation,
XX and/or function of cells that appear in the bone marrow, and leukocytes,
XX treat bone marrow, blood and hematopoietic associated diseases and
XX disorders, atelectasis, pulmonary congestion or oedema, emphysema,
XX chronic bronchitis, bronchial asthma and bronchiectasis, intestinal
XX disorders, spleen associated diseases, modulate renal disorders, treat
XX cardiovascular disorders such as ischemic heart disease, modulate the
XX proliferation, differentiation, and/or function of bone and cartilage
XX cells and to treat bone and/or cartilage associated diseases or
XX disorder. They may also be used to treat disorders associated with the
XX ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic
XX disease, inflammations, bacterial and viral meningitis, Alzheimer's
XX disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,
XX brain cancers, hydrocephalus and encephalitis, and treat hepatic
XX disorders
```

```
XX SQ Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 other;
Query Match 90.6%; Score 961; DB 21; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 ATGCGGAGCTGGAGTTTGTTCACATCATCATCATCTGTTGGTGTGATGATGTTGATGTTG 160
DB 6 ATGCGGAGCTGGAGTTTGTTCACATCATCATCATCTGTTGGTGTGATGATGTTGATGTTG 65
QY 161 GTGTGTATCAGTGCCTGCTGAGCCACTTACAAAGTGTCTGCACGGTCTCTTCATCAGCCGG 220
DB 66 GTGTGTATCAGTGCCTGCTGAGCCACTTACAAAGTGTCTGCACGGTCTCTTCATCAGCCGG 125
QY 221 CACAGCAGGGGGCGAGAGAGAGATGCTTCTCTCAGAAAGATGCTGTGGCCCTCG 280
DB 126 CACAGCAGGGGGCGAGAGAGATGCTTCTCTCAGAAAGATGCTGTGGCCCTCG 185
QY 281 GAGAGCAGTGTCTAGCAACGGAATCCAGAGCCGAGGTCTACGCCGCCCTCGGCC 340
DB 186 GAGAGCAGTGTCTAGCAACGGAATCCAGAGCCGAGGTCTACGCCGCCCTCGGCC 245
QY 341 ACCGACCGCTGGCGCTGCTGCCCGCTTCCGCCAGGGGAGCGCTTCCACCGCTTCCAGCCC 400
DB 246 ACCGACCGCTGGCGCTGCTGCCCGCTTCCGCCAGGGGAGCGCTTCCACCGCTTCCAGCCC 305
QY 401 ACCTATCGTACCTGTCAGCAGAGATGACCTGCGGCCACCATCTCGCTGTACAGCGGG 460
DB 306 ACCTATCGTACCTGTCAGCAGAGATGACCTGCGGCCACCATCTCGCTGTACAGCGGG 365
QY 461 GAGGAGCCCCCACCCTTACCAGGGCCCTGTCAGCCCTTCAGCTTCGGGAGCCCGAGCAGCAG 520
DB 366 GAGGAGCCCCCACCCTTACCAGGGCCCTGTCAGCCCTTCAGCTTCGGGAGCCCGAGCAGCAG 425
QY 521 CTGGAATCTGAACCGGGAGTCTGGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGAC 580
DB 426 CTGGAATCTGAACCGGGAGTCTGGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGAC 485
QY 581 CTGATGATAGTGTCCAGGCTGGGGGGCCCTTCCGCCAGCGAGTAACCTCGGGCATCAGC 640
DB 486 CTGATGATAGTGTCCAGGCTGGGGGGCCCTTCCGCCAGCGAGTAACCTCGGGCATCAGC 545
QY 641 GCCACGTGTACGCGCGGCGCGCATGAGGGGGCGCCGCCACCTACAGCGAGGTC 700
DB 546 GCCACGTGTACGCGCGGCGCGCATGAGGGGGCGCCGCCACCTACAGCGAGGTC 605
QY 701 ATCGGCCCACTACCGGGGTCTCTCTCCAGCACCCAGCAGAGTGGGGCCGCCCTCTCTTG 760
DB 606 ATCGGCCCACTACCGGGGTCTCTCTCCAGCACCCAGCAGAGTGGGGCCGCCCTCTCTTG 665
QY 761 CTGAGAGGGGACCCGGCTCCACACACATCGGGCCCTTAGAGAGCGCAGCCATCTGG 820
DB 666 CTGAGAGGGGACCCGGCTCCACACACATCGGGCCCTTAGAGAGCGCAGCCATCTGG 725
QY 821 AGCAAGAGAGAGGATAAACAAGGACACCTCTTAGGGTCCCGGCGGGCGCGCT 880
DB 726 AGCAAGAGAGAGGATAAACAAGGACACCTCTTAGGGTCCCGGCGGGCGCGCT 785
QY 881 GGGGCTGCTAGGTGAAAAGGACAGAACTCCGGCTCTTTAGAGAGAGGAGTACAGAGAA 940
DB 786 GGGGCTGCTAGGTGAAAAGGACAGAACTCCGGCTCTTTAGAGAGAGGAGTACAGAGAA 845
QY 941 GCGGGGGGGCGCAGCAAGCATCTGTGGCCCTTCCCTCCACCTCCCTCTGTGTATAATA 1000
DB 846 GCGGGGGGGCGCAGCAAGCATCTGTGGCCCTTCCCTCCACCTCCCTCTGTGTATAATA 905
QY 1001 TTTACATGTGATGCTGTGCTGAATGCAACAGCTTAAGAGAGCTTGCAAAAAA 1060
DB 906 TTTACATGTGATGCTGTGCTGAATGCAACAGCTTAAGAGAGCTTGCAAAAAA 965
QY 1061 A 1061
```

Db 966 A 966

RESULT 4  
AAA75163  
ID AAA75163 standard; cDNA; 969 BP.  
XX AC AAA75163;  
XX DT 15-JAN-2001 (first entry)  
XX DE cDNA clone encoding a human TANGO 261 polypeptide.  
XX  
XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH CDS 6..764  
FT FT /\*tag= a  
FT FT /product= "TANGO 261"  
XX  
XX WO200052022-A1.  
XX  
XX 08-SEP-2000.  
XX  
XX 01-MAR-2000; 200WO-US05226.  
XX  
XX 01-MAR-1999; 99US-0122458.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
XX WPI; 2000-579269/54.  
XX P-PSDB; AAB18461.  
XX  
XX Novel human and murine secreted proteins designated TANGO 216, 261,  
XX 262, 266 and 267 useful as modulating agents of cellular processes,  
XX e.g. for treating cancer -  
XX  
XX Disclosure; Page -; 175pp; English.  
XX  
XX AAA75163-65 encode human TANGO 261 proteins. The specification also  
XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
XX polypeptides can be used to modulate cellular proliferation, modulate  
XX cellular differentiation and/or modulate cellular adhesion. The  
XX proteins can be used to treat any von Willebrand factor-associated  
XX disorder, regulate extracellular matrix structuring, cellular adhesion,  
XX and cell trafficking and/or migration, modulate cellular interactions,  
XX modulate cell adhesion in proliferative disorders, such as cancer,  
XX modulate the proliferation, differentiation, and/or function of cells  
XX that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
XX and hematopoietic associated diseases and disorders, atelectasis,  
XX pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
XX asthma and bronchiectasis, intestinal disorders, spleen associated  
XX diseases, modulate renal disorders, treat cardiovascular disorders such  
XX as ischemic heart disease, modulate the proliferation, differentiation,  
XX and/or function of bone and cartilage cells and to treat bone and/or  
XX cartilage associated diseases or disorder. They may also be used to  
XX treat disorders associated with the ovaries, and cerebral oedema.  
XX  
XX hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
XX bacterial and viral meningitis, Alzheimer's disease, cerebral  
XX toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers.  
XX

```
RESULT 5
ABK92120
ID ABK92120 standard; DNA; 1140 BP.
XX AC ABK92120;
XX DT 15-AUG-2002 (first entry)
XX DE Prostate cancer-associated DNA sequence #6.
XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX KW gene therapy; gene; ds.
XX OS Mammalia.
XX PN WO200230268-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US32045.
XX PR 13-OCT-2000; 2000US-0687576.
XX PR 08-DEC-2000; 2000US-0733288.
XX PR 08-DEC-2000; 2000US-0733742.
XX PR 24-JAN-2001; 2001US-263957P.
XX PR 16-MAR-2001; 2001US-276791P.
XX PR 16-MAR-2001; 2001US-276888P.
XX PR 06-APR-2001; 2001US-281922P.
XX PR 24-APR-2001; 2001US-286214P.
XX PR 30-APR-2001; 2001US-0847046.
XX PR 04-MAY-2001; 2001US-288589P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX DR P-PSDB; ABG61805.
XX PT Detecting a prostate cancer-associated transcript in a cell in a
XX PT patient, useful for diagnosing prostate cancer (PC) or screening
XX PT modulators of PC, by determining if prostate cancer-associated genes
XX PT are expressed in a prostate tissue
XX PS Claim 22; Page 305; 436pp; English.
XX CC The present invention relates to methods of detecting a prostate
XX CC cancer-associated transcript in a cell from a patient. The method
XX CC comprises contacting a biological sample from the patient with
XX CC prostate cancer-associated polynucleotides (designated PC genes) that
XX CC selectively hybridise to a sequence that is at least 80% identical
XX CC to them. The prostate cancer-associated polynucleotide sequences
XX CC are differentially expressed in prostate tumour tissue or in
XX CC prostate cancer and are derived from the tissues of various
XX CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
XX CC The methods of the invention are useful for diagnosing and treating
XX CC prostate cancer in mammals. The prostate cancer-associated genes are
XX CC useful for diagnosing or treating prostate cancer, as well as for
XX CC identifying modulators of prostate cancer or agents that inhibit
XX CC prostate cancer. The nucleic acid sequences are particularly useful
XX CC in gene therapy, as a vaccine or in antisense applications.
XX CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX CC sequences.
XX CC Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;
XX CC Query Match 87.3%; Score 926; DB 24; Length 1140;
XX CC Best Local Similarity 99.9%; Pred. NO. 0;
XX CC Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 81 GCTCTCTCGGAACACGAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGG 140
```

```
Db 76 GTCTCTCTCGGAACACGAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGG 135
QY 141 TGGTGATGATGGTGATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 200
Db 136 TGGTGATGATGGTGATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 195
QY 201 CACGGTCTCTTCATCAGCCGGCACACGAGGCGGAGGAGAGAGAGATGCCCTGTCTCAG 260
Db 196 CACGGTCTCTTCATCAGCCGGCACACGAGGCGGAGGAGAGAGAGATGCCCTGTCTCAG 255
QY 261 AAGGATGCTCTGCTGGCCCTTCGAGAGACACAGTGTTCAGGCAACGGAATCCAGAGCCGAGG 320
Db 256 AAGGATGCTCTGCTGGCCCTTCGAGAGACACAGTGTTCAGGCAACGGAATCCAGAGCCGAGG 315
QY 321 TCTACGCCCGCCCTCGGCCCAACGACCGCTTGGCGGTGGCGCCCTTCGCCCAGCGGGAGC 380
Db 316 TCTACGCCCGCCCTCGGCCCAACGACCGCTTGGCGGTGGCGCCCTTCGCCCAGCGGGAGC 375
QY 381 GCTTCCACCGCTTCAGGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCCA 440
Db 376 GCTTCCACCGCTTCAGGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCCA 435
QY 441 CCATCTCGCTGTACAGCGGGAGGAGCCGCCACCTACAGGGGCCCTTCGACCTCCAGC 500
Db 436 CCATCTCGCTGTACAGCGGGAGGAGCCGCCACCTACAGGGGCCCTTCGACCTCCAGC 495
QY 501 TTCGGGAGCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCTGGTGGCGGACCCGCCAACA 560
Db 496 TTCGGGAGCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCTGGTGGCGGACCCGCCAACA 555
QY 561 GAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
Db 556 GAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
QY 621 GCAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGGGGGCGCATGAGAGGGCGCGC 680
Db 616 GCAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGGGGGCGCATGAGAGGGCGCGC 675
QY 681 CGCCACCTACAGCGAGGTCTATCGGCCACTACCGGGGTCTCTTCCAGACACGACGAGA 740
Db 676 CGCCACCTACAGCGAGGTCTATCGGCCACTACCGGGGTCTCTTCCAGACACGACGAGA 735
QY 741 GCAGTGGCGCCCTCTCTTCTGCGAGGGGAGCCCGCTCCACACACACACATCGCGCCCC 800
Db 736 GCAGTGGCGCCCTCTCTTCTGCGAGGGGAGCCCGCTCCACACACACACATCGCGCCCC 795
QY 801 TAGAGAGCGCAGCCATCTTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTTAGG 860
Db 796 TAGAGAGCGCAGCCATCTTGGAGCAAGAGAGAGATTAACAGAAAGGACACCCCTCTTAGG 855
QY 861 GTCCCGAGGGGGCGCGGCTGGGGCTGGTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 920
Db 856 GTCCCGAGGGGGCGCGGCTGGGGCTGGTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 915
QY 921 TAGAGAGAGGTGAGAGGAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCC 980
Db 916 TAGAGAGAGGTGAGAGGAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCC 975
QY 981 CACCTCCCTCTGTATATAATTTATCATGTGATGTCTGGTCTGTAATGCAAGCTAAGAGA 1040
Db 976 CACCTCCCTCTGTATATAATTTATCATGTGATGTCTGGTCTGTAATGCAAGCTAAGAGA 1035
QY 1041 GCTTGCACAAAAA 1057
Db 1036 GCTTGCACAAAAA 1052
```

```
RESULT 6
AAA75164
ID AAA75164 standard; cDNA; 969 BP.
XX AC AAA75164;
```

XX	15-JAN-2001	(first entry)
DT	cDNA clone encoding a human TANGO 261 polypeptide.	
XX		
DE	TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;	
XX	cellular proliferation; cellular differentiation; cellular adhesion;	
KW	von Willebrand factor-associated disorder; cell trafficking; cancer;	
KW	hematopoietic associated disease; atelectasis; pulmonary congestion;	
KW	oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;	
KW	intestinal disorder; spleen associated disease; renal disorder;	
KW	cardiovascular disorder; ischemic heart disease; hydrocephalus;	
KW	brain herniation; iatrogenic disease; inflammation; meningitis;	
KW	Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;	
KW	multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	
FT	Location/Qualifiers	
FT	6..764	
FT	/*tag= a	
FT	/product= "TANGO 261"	
XX		
XX	W0200052022-A1.	
PN		
XX		
PD	08-SEP-2000.	
XX		
XX	01-MAR-2000; 2000WO-US05226.	
PF		
XX		
PR	01-MAR-1999; 99US-0122458.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
XX		
PI	Barnes TM, Holtzman DA, Sharp JD, Frazer CC;	
XX		
DR	WPI; 2000-579269/54.	
DR	P-PSDB; AAB18462.	
XX		
PT	Novel human and murine secreted proteins designated TANGO 216, 261,	
PT	262, 266 and 267 useful as modulating agents of cellular processes,	
PT	e.g. for treating cancer -	
XX		
PS	Disclosure; Page -: 175pp; English.	
XX		
CC	AAA75163-65 encode human TANGO 261 proteins. The specification also	
CC	describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO	
CC	polypeptides can be used to modulate cellular proliferation, modulate	
CC	cellular differentiation and/or modulate cellular adhesion. The	
CC	proteins can be used to treat any von Willebrand factor-associated	
CC	disorder, regulate extracellular matrix structuring, cellular adhesion,	
CC	and cell trafficking and/or migration, modulate cellular interactions,	
CC	modulate cell adhesion in proliferative disorders, such as cancer,	
CC	modulate the proliferation, differentiation, and/or function of cells	
CC	that appear in the bone marrow, and leukocytes, treat bone marrow, blood	
CC	and hematopoietic associated diseases and disorders, atelectasis,	
CC	pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial	
CC	asthma and bronchiectasis, intestinal disorders, spleen associated	
CC	diseases, modulate renal disorders, treat cardiovascular disorders such	
CC	as ischemic heart disease, modulate the proliferation, differentiation,	
CC	and/or function of bone and cartilage cells and to treat bone and/or	
CC	cartilage associated diseases or disorder. They may also be used to	
CC	treat disorders associated with the ovaries, and cerebral oedema,	
CC	hydrocephalus, brain herniations, iatrogenic disease, inflammations,	
CC	bacterial and viral meningitis, Alzheimer's Disease, cerebral	
CC	toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,	
CC	hydrocephalus and encephalitis, and treat hepatic disorders.	
CC	note: the present sequence does not appear in the specification; it was	
CC	created using information provided.	
XX		
SQ	Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;	
Query Match 85.8%; Score 910; DB 21; Length 969;		
Best Local Similarity 99.9%; Pred No. 0;		



	AAI57868	standard; cdNA; 1066 BP.
XX	AAI57868;	
DT	22-OCT-2001	(first entry)
XX	Human polynucleotide SEQ ID NO 71.	
XX	Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Draeger Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia; ss.	
OS	Homo sapiens.	
XX	WO200153312-A1.	
PN	26-JUL-2001.	
PD	26-DEC-2000; 2000WO-US34263.	
PF	21-JAN-2000; 2000US-0488725.	
XX	PR 25-APR-2000; 2000US-0552317.	
FR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	18-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX	(HYSE-) HYSEQ INC.	
PA	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
FI	Zhao QA, Zhou P, Goodrich R, Drimanac RT;	
XX	WPI; 2001-442253/47.	
DR	P-PSDB; AAM38712.	
XX	Novel nucleic acids and polypeptides, useful for treating disorders	
PT	such as central nervous system injuries -	
FT	Claim 1; SEQ ID NO 71; 10078pp; English.	
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotropic,	
CC	immunosuppressant and cytosolic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathies and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Draeger Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokine activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX	Sequence 1066 BP; 225 A; 348 C; 325 G; 168 T; 0 other;	
SQ	Query Match 84.0%; Score 891; DB 22; Length 1066;	
	Best Local Similarity 99.9%; Pred.No.; Indels 0; Mismatches 0;	
	Mismatches 941; Conservative 0; Matches 1; Indels 0; Gaps 0;	
QY	117 TTGTTTCAGATCATCATCGTGGTGATGATGTGGTGATGCACGTGCC 176	



XX Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;  
 KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;  
 KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
 KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
 KW heart failure.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 413..1276  
 FT /\*tag= a  
 FT /product= "MIVR-1"  
 FT /note= "This region is specifically claimed in claim 3"  
 XX  
 PN WO200216416-A2.  
 XX  
 XX 28-FEB-2002.  
 XX  
 XX 21-AUG-2001; 2001WO-US26089.  
 XX  
 XX 22-AUG-2000; 2000US-227159P.  
 XX  
 XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 XX (PFIZ ) PFIZER INC.  
 XX  
 XX Lee RT, Landeschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 XX  
 XX WPI; 2002-280912/32.  
 DR P-PSDB; AAU78231.  
 XX  
 XX Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 PT Receptor-1 polypeptide, useful for treating cardiovascular diseases  
 XX  
 XX Claim 2; Page 87-88; 105pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences AI761441.1, AI594390,  
 CC NM 004338 and A0177461. Also included are expression vectors, host  
 CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
 CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
 CC with a candidate agent, where the molecule is a nucleic acid molecule  
 CC comprising MIVR-1, IEX-1, VDRP-1, BIG-2 and FIS-11d or its  
 CC expression product, determining if the anti-apoptotic activity is  
 CC modulated and thereby identifying a modulator. The cardiac cell anti-  
 CC apoptotic molecules and nucleic acids of the invention are useful for  
 CC treating, diagnosing and monitoring progression of such diseases and  
 CC disorders as characterised by increased apoptotic cell-death of vascular  
 CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
 CC stroke, arteriosclerosis and heart failure. The present sequence  
 CC encodes human MIVR-1.  
 XX  
 XX Sequence 1321 BP; 223 A; 493 C; 440 G; 165 T; 0 other;  
 SQ

Query Match 70.6%; Score 749; DB 24; Length 1321;  
 Best Local Similarity .99.9%; Pred. No. 1.5e-289;  
 Matches 799; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 105 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGGTGGTGG 164  
 DB 522 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGGTGGTGG 581  
 QY 165 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 224  
 DB 582 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 641  
 QY 225 GCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGGATGCCCTGTGGCCCTCGGAGA 284  
 DB 642 GCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGGATGCCCTGTGGCCCTCGGAGA 701  
 QY 285 GCACAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCACCG 344

DB 702 GCACAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCCACCG 761  
 QY 345 ACCGCTGGCGTGTGCGGCTTTCGCCAGCGGGAGCGTTCACCGCTTTCAGCCACCT 404  
 DB 762 ACCGCTGGCGTGTGCGGCTTTCGCCAGCGGGAGCGTTCACCGCTTTCAGCCACCT 821  
 QY 405 ATCCGTACTCTGACGACGACGATCGACCTTGCCTGCCGCCACCATCTCGCTGTGACGCGGAGG 464  
 DB 822 ATCCGTACTCTGACGACGACGATCGACCTTGCCTGCCGCCACCATCTCTCGTGTGACGCGGAGG 881  
 QY 465 AGCCCCCACCCTTACCAGGGCCCTGCACCTCTCAGCTTGGGACCCCGAGCAGCAGCTGG 524  
 DB 882 AGCCCCCACCCTTACCAGGGCCCTGCACCTCTCAGCTTGGGACCCCGAGCAGCAGCTGG 941  
 QY 525 AACTGAACCGGGAGTCCGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGA 584  
 DB 942 AACTGAACCGGGAGTCCGTGGCGCACCCCCCAACAGAACCATCTTCGACAGTGA 1001  
 QY 585 TGGATAGTCCAGGCTGGCGGCCCTTGCCTGCCGCCACCATCTTCGACAGTGA 644  
 DB 1002 TGGATAGTCCAGGCTGGCGGCCCTTGCCTGCCGCCACCATCTTCGACAGTGA 1061  
 QY 645 CGTGTACTGGCAGCGCGCGGCATGAGAGGGCCCGCCGACCTTACAGCAGGTCATCG 704  
 DB 1062 CGTGTACTGGCAGCGCGCGGCATGAGAGGGCCCGCCGACCTTACAGCAGGTCATCG 1121  
 QY 705 GCCACTACCCGGGTCTCTCTTCCAGCACACACATCGCGCCCTTACAGCAGGTCATCG 764  
 DB 1122 GCCACTACCCGGGTCTCTCTTCCAGCACACACATCGCGCCCTTACAGCAGGTCATCG 1181  
 QY 765 AGGGACCCGGTCTCCACACACACATCGCGCCCTTACAGCAGGTCATCG 824  
 DB 1182 AGGGACCCGGTCTCCACACACACATCGCGCCCTTACAGCAGGTCATCG 1241  
 QY 825 AAGAGAAGGATAAAGAGAGACACCTCTCTAGGTCCTCCAGGGGGCGCGGCTGGG 884  
 DB 1242 AAGAGAAGGATAAAGAGAGACACCTCTCTAGGTCCTCCAGGGGGCGCGGCTGGG 1301  
 QY 885 CTGCGTAGGTGAAGAGGAG 904  
 DB 1302 CTGCGTAGGTGAAGAGGAG 1321  
 RESULT 10  
 ID AAI59654/c  
 ID AAI59654 standard; cDNA; 1069 BP.  
 XX AC AAI59654;  
 XX AC AAI59654;  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 3643.  
 XX  
 KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200153312-A1.  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.





or comprising an amino acid sequence encoded by one of 136 nucleotide  
prostate specific nucleic acids sequences, PSNA, (or a sequence that  
hybridizes to it or is 60% identical to it), given in the specification.  
Also included are a vector comprising the polynucleotide, a host cell  
comprising the vector, an antibody specific for the PSP proteins and a  
vaccine comprising the protein or polynucleotide. The PSP and PSNA are  
useful for diagnosing and monitoring the presence and metastases of  
prostate cancer in a patient. The PSNA is useful for determining the  
level of PSNA in a sample. An antibody to the PSP is useful for  
determining the presence of prostate specific protein in a sample, and  
for treating a patient with prostate cancer, which induces an immune  
response against the prostate cancer cell expressing the nucleic acid or  
polypeptide and a kit is useful for detecting a risk of cancer or  
presence of cancer in a patient. PSNA is useful as hybridisation probes  
to detect, characterise and quantify hybridising nucleic acids from both  
genomic and transcript-derived nucleic acid samples and also in  
microarrays. Sequences of PSP and PSNA are useful as components in  
databases for search analysis as well as in sequence analysis algorithms.  
PSNA is useful to drive in vivo expression of PSP. The present  
sequence is a PSNA of the invention.

Sequence 1583 BP; 361 A; 469 C; 457 G; 296 T; 0 other;

Query Match 48.0%; Score 509; DB 24; Length 1583;  
Best Local Similarity 99.8%; Pred. No. 9.9e-194;  
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 313 GCGCAGAGTCTACGCGCCCTCGGCCACCGACCGCTGGCGCGCCCTTCGCCCA 372  
DB 827 GCGCAGAGTCTACGCGCCCTCGGCCACCGACCGCTGGCGCGCCCTTCGCCCA 886  
QY 373 GCGGAGCGCTTCCACCGCTTCCAGCCACCTATCGTACCTGCAGCAGCAGATGACCT 432  
DB 887 GCGGAGCGCTTCCACCGCTTCCAGCCACCTATCGTACCTGCAGCAGCAGATGACCT 946  
QY 433 GCGGCGCCACATCTCGCTGTACAGCGGGAGGAGCGCCACCTACCGAGGCGCCCTGCAC 492  
DB 947 GCGGCGCCACATCTCGCTGTACAGCGGGAGGAGCGCCACCTACCGAGGCGCCCTGCAC 1006  
QY 493 CTCTCAGCTTCGGGACCGCGAGCAGCAGTGGAACTGAAACCGGGAGTGGTGGCGGACC 552  
DB 1007 CTCTCAGCTTCGGGACCGCGAGCAGCAGTGGAACTGAAACCGGGAGTGGTGGCGGACC 1066  
QY 553 CCCAACAGAACATCTTCGACAGTGCACCTGTATGTATAGTCCAGGCTGGGCGCCCTG 612  
DB 1067 CCCAACAGAACATCTTCGACAGTGCACCTGTATGTATAGTCCAGGCTGGGCGCCCTG 1126  
QY 613 CCCCCCAGCAGTAATCTCGGCGATCAGCGCCACGTCTACGCGAGCGGGCGCGCATGGA 672  
DB 1127 CCCCCCAGCAGTAATCTCGGCGATCAGCGCCACGTCTACGCGAGCGGGCGCGCATGGA 1186  
QY 673 GGGGCGCGCGCCACCTACAGGAGGTATCGGCACTACCGGGGTCTCTTCACGA 732  
DB 1187 GGGGCGCGCGCCACCTACAGGAGGTATCGGCACTACCGGGGTCTCTTCACGA 1246  
QY 733 CCAGCAGCAGCTGGGCGCCCTCTCTGTGGAGGGACCGGCTCCACACACACAT 792  
DB 1247 CCAGCAGCAGCTGGGCGCCCTCTCTGTGGAGGGACCGGCTCCACACACACAT 1306  
QY 793 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACC 852  
DB 1307 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATTAACAGAAAGGACACC 1366  
QY 853 TCTTAGGTTCCTCCAGGGGG 872  
DB 1367 TCTTAGGTTCCTCCAGGGGG 1386

RESULT 12  
AB284732/c  
ID AB284732 standard; cDNA; 474 BP.  
XX  
AC AB284732;

XX 14-MAY-2003 (first entry)  
DT Toxicologically relevant human nucleotide sequence #1891.  
DE Toxicologically relevant gene; toxicological response; gene; ss.  
KW Homo sapiens.  
OS WO2003016500-A2.  
PD 27-FEB-2003.  
PF 16-AUG-2002; 2002WO-US26514.  
PR 16-AUG-2001; 2001US-313080P.  
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.  
XX Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schweiser K;  
PI Alen P;  
XX WPI; 2003-268322/26.  
DR Determining a toxicological response to an agent, useful for screening  
of drugs, comprises comparing the expression profile of one or more  
human toxic response genes to a reference gene expression profile  
indicative of toxicity -  
PS Claim 1; Page 444; 455pp; English.  
CC The present invention describes a method (M1) for determining a  
toxicological response to an agent, which comprises comparing the  
expression profile of one or more human toxic response genes to a  
reference gene expression profile indicative of toxicity, and so  
determining the presence of a toxic response to the agent. Also  
described: (1) an array comprising one or more polynucleotides selected  
from the genes corresponding to the partial sequences given in AB282842  
to AB284744, or their fragments of at least 20 nucleotides, or  
homologues; and (2) determining if a gene putatively identified to be a  
toxic response gene plays a role on toxic response pathways by  
determining the expression profile of the gene after exposure of cells  
or a human subject to a known toxic pharmaceutical or industrial agent,  
comprising: (a) exposing cells to an agent or isolating cells from a  
human subject who was exposed to an agent; (b) obtaining the test gene  
expression profile for a putatively identified toxic response gene after  
exposure to a known toxic pharmaceutical or industrial agent; and  
(c) comparing the test profile to the expression profile of a gene with  
a similar function or comparing the test profile to the expression  
profile of that gene after exposure to other known toxic compounds. The  
methods are useful for predicting and determining toxicological responses  
on a cellular, organ or system level. The arrays comprising the human  
genes are useful for toxicological screening of drugs, pharmaceutical  
compounds and chemicals.

Sequence 474 BP; 62 A; 136 C; 186 G; 90 T; 0 other;

Query Match 39.1%; Score 415; DB 25; Length 474;  
Best Local Similarity 100.0%; Pred. No. 3.9e-156;  
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 297 GCACGGGAATCCAGAGCGCGAGGTCTACGCCCGCTCGGCCACCGCCCTGGCGG 356  
DB 474 GCACGGGAATCCAGAGCGCGAGGTCTACGCCCGCTCGGCCACCGCCCTGGCGG 415  
QY 357 TGCGGCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCAGGCCACCTATCGTACCTGC 416  
DB 414 TGCGGCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCAGGCCACCTATCGTACCTGC 355  
QY 417 AGCAGGAGATGACCTGCGGCCCGCCACCATCTCGCTGTACAGCGGAGGAGCCCGCCCT 476  
DB 354 AGCAGGAGATGACCTGCGGCCCGCCACCATCTCGCTGTACAGCGGAGGAGCCCGCCCT 295



CC compositions for treatment of pancreatic tumors. AAZ52858-253014  
CC represent expressed sequence tag (EST) fragments derived from a human  
CC pancreatic tumor cDNA library and which encode the proteins represented  
CC in AAY73814-Y74252.

XX Sequence 812 BP; 157 A; 272 C; 237 G; 146 T; 0 other;

Query Match 30.3%; Score 322; DB 20; Length 812;  
Best Local Similarity 100.0%; Pred. No. 5e-119; Mismatches 0; Indels 0; Gaps 0;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 TCACGTGCTGCTGAGCCACTAAGCTGTCTGCAGGTCCTTCATCAGCGGCACAGCC 227  
DB |||||  
QY 426 TCACGTGCTGCTGAGCCACTAAGCTGTCTGCAGGTCCTTCATCAGCGGCACAGCC 485  
DB |||||  
QY 228 AGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGAGCA 287  
DB |||||  
QY 486 AGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGAGCA 545  
DB |||||  
QY 288 CAGTGTGAGGCAACGGAATCCAGAGCGCGAGGTCTAGCCCGCTCGGCCACCGACC 347  
DB |||||  
QY 546 CAGTGTGAGGCAACGGAATCCAGAGCGCGAGGTCTAGCCCGCTCGGCCACCGACC 605  
DB |||||  
QY 348 GCCTGGCGGTGCGGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACTATC 407  
DB |||||  
QY 606 GCCTGGCGGTGCGGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCAGGCCACTATC 665  
DB |||||  
QY 408 CGTACTGACGACGAGATCGACCTGCGCGCCACCATCTCGCTGTCAGAGGGAGGAGC 467  
DB |||||  
QY 666 CGTACTGACGACGAGATCGACCTGCGCGCCACCATCTCGCTGTCAGAGGGAGGAGC 725  
DB |||||  
QY 468 CCCACCCCTTACCAGGCGCCCTG 489  
DB |||||  
QY 726 CCCACCCCTTACCAGGCGCCCTG 747  
DB |||||

RESULT 15  
ABK12143/c  
ID ABK12143 standard; cDNA; 693 BP.  
XX  
AC ABK12143;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Human MIVR-1 homologous sequence #1.  
XX  
KW Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;  
KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;  
KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
KW heart failure; A1761441.1.  
XX  
OS Homo sapiens.  
XX  
PN WO200216416-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 21-AUG-2001; 2001WO-US26089.  
XX  
PR 22-AUG-2000; 2000US-227159P.  
XX  
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.  
PA (PFIZ) PFIZER INC.  
XX  
PI Lee RT, Landeschulz KT, Kennedy SP, Thompson JF, Turi TG;  
XX  
DR WPI; 2002-280912/32.  
XX  
PT Novel nucleic acid molecule encoding Mechanically Induced Vascular  
PT Receptor-1 polypeptide, useful for treating cardiovascular diseases  
XX  
PS Disclosure, Page 101; 105pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a  
CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
CC cardiac cell anti-apoptotic activity and fragments of it provided  
CC they are not identical to Genbank sequences A1761441.1, A1594390,  
CC NM\_004338 and A0177461. Also included are expression vectors, host  
CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of  
CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity  
CC with a candidate agent, where the molecule is a nucleic acid molecule  
CC comprising MIVR-1, IEX-1, VDUP-1, BTG-2 and TIS-11d or its  
CC expression product, determining if the anti-apoptotic activity is  
CC modulated and thereby identifying a modulator. The cardiac cell anti-  
CC apoptotic molecules and nucleic acids of the invention are useful for  
CC treating, diagnosing and monitoring progression of such diseases and  
CC disorders as characterised by increased apoptotic cell death of vascular  
CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,  
CC stroke, arteriosclerosis and heart failure. The present sequence  
CC is one of the four Genbank sequences (A1761441.1) which are homologous to  
CC the cDNA for human MIVR-1 and which are specifically disclaimed.

XX Sequence 693 BP; 101 A; 205 C; 237 G; 149 T; 1 other;

Query Match 29.5%; Score 313; DB 24; Length 693;  
Best Local Similarity 99.3%; Pred. No. 2e-115;  
Matches 583; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 468 CCCACCCCTTACCAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAAC 527  
DB |||||  
QY 587 CCCACCCCTTACCAGGCGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAAC 528  
DB |||||  
QY 528 TGAACCGGGAGTCCGTGCGCGCACCCCAACAGAACCATCTTCGACAGTCACTGATGG 587  
DB |||||  
QY 527 TGAACCGGGAGTCCGTGCGCGCACCCCAACAGAACCATCTTCGACAGTCACTGATGG 468  
DB |||||  
QY 588 ATAGTGCCAGGCTGGGGCGCCCTGCGCCCGCCAGCAGTAATCGGGCATCAGCGCACGT 647  
DB |||||  
QY 467 ATAGTGCCAGGCTGGGGCGCCCTGCGCCCGCCAGCAGTAATCGGGCATCAGCGCACGT 408  
DB |||||  
QY 648 GCTACGGCAGCGCGGGCGGCATGAGAGGGCGCGCCCGCCACCTACAGCAGGTCACTGCGCC 707  
DB |||||  
QY 407 GCTACGGCAGCGCGGGCGGCATGAGAGGGCGCGCCCGCCACCTACAGCAGGTCACTGCGCC 348  
DB |||||  
QY 708 ACTACCGGGGTCTCTCTTCCAGCACCGAGCAGTGGGGCGCCCTCTTGTCTGGA-G 766  
DB |||||  
QY 347 ACTACCGGGGTCTCTCTTCCAGCACCGAGCAGTGGGGCGCCCTCTTGTCTGGA-G 288  
DB |||||  
QY 767 GGGACCGGGTCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAA 826  
DB |||||  
QY 287 GGGACCGGGTCTCCCGCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAA 228  
DB |||||  
QY 827 GAGAAGGNTAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCT 886  
DB |||||  
QY 227 GAGAAGGNTAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCT 168  
DB |||||  
QY 887 GCGTAGGTGAAAAGGCGAGAACACTCCCGCTCTTCTAGAGAGGAGTGAAGAAAGGCGGG 946  
DB |||||  
QY 167 GCGTAGGTGAAAAGGCGAGAACACTCCCGCTCTTCTAGAGAGGAGTGAAGAAAGGCGGG 108  
DB |||||  
QY 947 GGGCGCAGCAACGCATCGTGTGGCCCTCCCGCTCCCGACCTCCCTGTGTATAAATTTTACA 1006  
DB |||||  
QY 107 GGGCGCAGCAACGCATCGTGTGGCCCTCCCGCTCCCGACCTCCCTGTGTATAAATTTTACA 48  
DB |||||  
QY 1007 TGTGATGCTGCTGTAATGCAACAGCTTAAGAGAGCTTGCAAAAAA 1053  
DB |||||  
QY 47 TGTGATGCTGCTGTAATGCAACAGCTTAAGAGAGCTTGCAAAAAA 1

Search completed: December 8, 2003, 04:24:34  
Job time : 348 secs





; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-5

Query Match 2.3%; Score 24; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 900 GGCAGACACTCCGGCTTCTTAG 923  
|||||  
Db 1 GGCAGACACTCCGGCTTCTTAG 24

RESULT 4  
US-09-769-482-6/c  
; Sequence 6, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-6

Query Match 2.3%; Score 24; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1022 GAATGCACAAGCTAAGAGAGCTTG 1045  
|||||  
Db 24 GAATGCACAAGCTAAGAGAGCTTG 1

RESULT 5  
US-09-769-482-10/c  
; Sequence 10, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-10

Query Match 2.3%; Score 24; DB 4; Length 24;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1022 GAATGCACAAGCTAAGAGAGCTTG 1045  
|||||  
Db 24 GAATGCACAAGCTAAGAGAGCTTG 1

RESULT 6  
US-08-332-766A-1  
; Sequence 1, Application US/08332766A  
; Patent No. 5843647  
; GENERAL INFORMATION:  
; APPLICANT: JEFFREYS, Alec J.  
; APPLICANT: ARMOUR, John  
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
; NUMBER OF SEQUENCES: 125  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,766A  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9326052.9  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIRD, Donald J.  
; REGISTRATION NUMBER: 25,323  
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-332-766A-1

Query Match 2.2%; Score 23; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TCGTGTGATGATGATGATGATG 160  
Db 123 TCGTGTGATGATGATGATGATG 145

## RESULT 7

US-09-769-482-7  
; Sequence 7, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-7

Query Match 2.0%; Score 21; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTTGGTTCGGGTGAAGGC 26  
Db 1 CTTGGTTCGGGTGAAGGC 21

## RESULT 8

US-08-522-421-7  
; Sequence 7, Application US/08522421  
; Patent No. 5908973  
; GENERAL INFORMATION:  
; APPLICANT: Abu-Bakar, Umi Kalsom  
; APPLICANT: Barton, Sarah Louise  
; APPLICANT: Gallego-Veigas, Pedro Pablo  
; APPLICANT: Gray, Julie Elizabeth  
; APPLICANT: Grierson, Donald  
; APPLICANT: Lowe, Alexandra Louise  
; APPLICANT: Picton, Steve  
; APPLICANT: Whotton, Lee Colin  
; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/522,421  
; FILING DATE: 11-JAN-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305868.3  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305869.1  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305859.2  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305865.9  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305866.7  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305867.5  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305860.0  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9305862.6  
; FILING DATE: 22-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9314351.9  
; FILING DATE: 12-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9320988.0  
; FILING DATE: 12-OCT-1993

; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 686 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; ORGANISM: ERT17  
US-08-522-421-7

Query Match 2.0%; Score 21; DB 2; Length 686;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1041 GCTTCGCAAAAAAAAAAAAAA 1061  
Db 663 GCTTCGCAAAAAAAAAAAAAA 683

## RESULT 9

US-08-165-315D-3  
; Sequence 3, Application US/08165315D  
; Patent No. 5525716  
; GENERAL INFORMATION:  
; APPLICANT: Odd-Arne Olsen  
; APPLICANT: Roger Kalla  
; TITLE OF INVENTION: Promoter  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: St. Onge, Steward, Johnston & Reens  
; STREET: 986 Bedford Street  
; CITY: Stamford  
; STATE: Connecticut  
; COUNTRY: U.S.A.  
; ZIP: 06905  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" 1.44 Mb diskette  
; COMPUTER: IBM PC



OPERATING SYSTEM: MS DOS  
SOFTWARE: Word Processor  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/165,315D  
FILING DATE: 10 December 1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9324707.0  
FILING DATE: 2 December 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary M. Krinsky  
REGISTRATION NUMBER: 32423  
REFERENCE/DOCKET NUMBER: 2105-P0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-324-6155  
TELEFAX: 201-327-1096  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1327  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FRAGMENT TYPE: gene  
FEATURE:  
NAME/KEY: ltp2 gene  
US-08-165-315D-3

Query Match 2.0%; Score 21; DB 1; Length 1327;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 GTGATGATGGTGGTGGTG 163  
|||||  
DB 918 GTGATGATGGTGGTGGTG 938

RESULT 10  
US-08-046-585-15/c  
Sequence 15, Application US/08046585  
Patent No. 545362  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
APPLICANT: Herr, Winship  
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/046,585  
FILING DATE: 12-APR-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57503-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 398-3249  
TELEFAX: 910 277299

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-046-585-15

Query Match 2.0%; Score 21; DB 1; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGGTGGTGGTGG 158  
|||||  
DB 2748 TGGTGGTGGTGGTGGTGG 2728

RESULT 11  
US-08-393-703-15/c  
Sequence 15, Application US/08393703  
Patent No. 5585239  
GENERAL INFORMATION:  
APPLICANT: Lamarco, Kelly  
APPLICANT: Wilson, Angus  
APPLICANT: Herr, Winship  
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,703  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-57503-2/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-393-703-15

Query Match 2.0%; Score 21; DB 1; Length 8252;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGGTGGTGGTGG 158  
|||||  
DB 2748 TGGTGGTGGTGGTGGTGG 2728

RESULT 12  
PCT-US93-11721-15/c

```
; Sequence 15, Application PC/TUS9311721
; GENERAL INFORMATION:
; APPLICANT: Lamarco, Kelly
; APPLICANT: Wilson, Angus
; APPLICANT: Herr, Winship
; TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
; TITLE OF INVENTION: HOST CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBERTITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11721
; FILING DATE: 03-DEC-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: FP-57503-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US93-11721-15

Query Match 2.0%; Score 21; DB 5; Length 8252;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGATGGTGATGG 158
      |||||
Db 2748 TGGTGGTGATGATGGTGATGG 2728

RESULT 13
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
```

NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature

LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 2.0%; Score 21; DB 4; Length 1664976;  
Best Local Similarity 100.0%; Pred. No. 6.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0

Qy 138 TGGTGGTGGTGGTGGTGG 158  
Db 568047 TGGTGGTGGTGGTGGTGG 568027

RESULT 14  
US-09-769-482-9  
Sequence 9, Application US/09769482  
Patent No. 6566130  
GENERAL INFORMATION:  
APPLICANT: SRIVASTAVA, SHIV  
APPLICANT: MOUL, JUDD W.  
APPLICANT: XU, LINDA L.  
APPLICANT: SEGAWA, TAKEHIKO  
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
FILE REFERENCE: 04995.0057-00000  
CURRENT APPLICATION NUMBER: US/09/769,482  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,772  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,045  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Primer  
US-09-769-482-9

Query Match 1.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 363 CCTTCGCCCGGAGGCGC 382  
Db 1 CCTTCGCCCGGAGGCGC 20

RESULT 15  
US-08-753-247-22/c  
Sequence 22, Application US/08753247  
Patent No. 6210929  
GENERAL INFORMATION:  
APPLICANT: SCHLOKAT, Uwe  
APPLICANT: FISCHER, Bernhard  
APPLICANT: FALKNER, Falko-Guenther  
APPLICANT: DORNER, Friedrich  
APPLICANT: EIBL, Johann  
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN  
TITLE OF INVENTION: DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A  
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500

```

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,247
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AT 1928/95
; FILING DATE: 24-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40433/149
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-753-247-22

```

```

Query Match      1.9%; Score 20; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      144  TGATGATGGTGGTGGTG 163
Db      42  TGATGATGGTGGTGGTG 23

```

```

Search completed: December 8, 2003, 06:19:08
Job time : 85 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 03:56:15 ; Search time 2638 Seconds  
(without alignments)  
9775.227 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	722	68.0	729	13	BQ575741
c 2	665	62.7	967	13	BQ641849
c 3	652	61.5	655	13	BQ691705
c 4	652	61.5	951	9	AL558881

5	609	57.4	609	13	BQ636742
6	601	56.6	890	13	BQ690750
c 7	599	56.5	602	14	CA431191
c 8	585	55.1	588	13	BQ624784
c 9	570	53.7	629	13	BQ730650
c 10	559	52.7	559	10	BE855409
c 11	556	52.4	626	12	BM974296
c 12	536	50.5	570	13	BQ575582
c 13	536	50.5	945	13	BQ539219
c 14	530	50.0	563	14	CB049800
c 15	528	49.8	844	13	BQ686793
c 16	528	49.8	952	13	BQ157959
c 17	523	49.3	730	12	BM677602
c 18	508	47.9	1046	12	BM922276
c 19	505	47.6	1046	13	BQ559841
c 20	499	47.0	551	12	BM141979
c 21	492	46.4	633	12	BM714472
c 22	491	46.3	782	12	BQ015170
c 23	484	45.6	547	12	BM676516
c 24	484	45.6	552	12	BM713900
c 25	480	45.2	728	13	BQ683523
c 26	476	44.9	1007	9	AL558882
c 27	469	44.2	613	10	BQ680325
c 28	468	44.1	964	13	BQ559860
c 29	461	43.4	461	12	BM712680
c 30	460	43.4	850	13	BQ602918
c 31	449	42.3	1280	13	BQ691500
c 32	443	41.8	446	12	BM681946
c 33	442	41.7	973	13	BQ169156
c 34	441	41.6	668	14	CB044866
c 35	441	41.6	916	13	BQ954555
c 36	440	41.5	938	13	BQ157842
c 37	436	41.1	437	9	A1936228
c 38	436	41.1	502	9	A1921394
c 39	411	38.7	618	14	CD367193
c 40	410	38.6	1127	13	BQ174654
c 41	402	37.9	1201	9	AL517150
c 42	399	37.6	451	9	A1493698
c 43	392	36.9	874	13	BX362396
c 44	388	36.6	588	9	A1377498
c 45	387	36.5	619	9	A1742327

#### ALIGNMENTS

RESULT 1  
BQ575741/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ575741 729 bp mRNA linear EST 19-JUN-2002  
UI-H-EZ1-bbg-h-14-0-UI-s1 NCI CGAP Ch2 Homo sapiens CDNA clone  
UI-H-EZ1-bbg-h-14-0-UI 3', mRNA sequence.

BQ575741 GI:21479058

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 729)

NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of

Orthopaedics

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)

Seq primer: M13 FORWARD



```
QY 227 CAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGAGAGC 286
|
|
|
Db 121 CAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGAGAGC 180
|
|
|
QY 287 ACAGTGTGAGGCAACGGAATCCAGAGCGCAGGTCTAGCGCCCGCTCGGCCACCGAC 346
|
|
|
Db 181 ACAGTGTGAGGCAACGGAATCCAGAGCGCAGGTCTAGCGCCCGCTCGGCCACCGAC 240
|
|
|
QY 347 CCGCTGGCGGTCCCGCTTCCAGAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTAT 406
|
|
|
Db 241 CCGCTGGCGGTCCCGCTTCCAGAGCGGAGCGCTTCCACCGCTTCCAGCGCCACCTAT 300
|
|
|
QY 407 CCGTACCTGACAGCAGATCGACCTGCGGCCCAACATCTCCTGTTCAGCGGGAGGAG 466
|
|
|
Db 301 CCGTACCTGACAGCAGATCGACCTGCGGCCCAACATCTCCTGTTCAGCGGGAGGAG 360
|
|
|
QY 467 CCGCCACCTACAGCGGCCCTTCCAGCTTCCAGCTTCCAGCGGCCCGAGCGAGTGA 526
|
|
|
Db 361 CCGCCACCTACAGCGGCCCTTCCAGCTTCCAGCTTCCAGCGGCCCGAGCGAGTGA 420
|
|
|
QY 527 CTGAACCGGAGTCTGCGCGCACCCCAACAGAACCATCTTCGACAGTCACTGTATG 586
|
|
|
Db 421 CTGAACCGGAGTCTGCGCGCACCCCAACAGAACCATCTTCGACAGTCACTGTATG 480
|
|
|
QY 587 GATAGTGCAGCTGGCGGCCCTTCCAGCGAGTAACTCGGGCATCAGCGCCACG 646
|
|
|
Db 481 GATAGTGCAGCTGGCGGCCCTTCCAGCGAGTAACTCGGGCATCAGCGCCACG 540
|
|
|
QY 647 TGCTACGCGAGCGGGCGGATGAGGGCGCGCCCACTCAGCGAGTCACTCGGC 706
|
|
|
Db 541 TGCTACGCGAGCGGGCGGATGAGGGCGCGCCCACTCAGCGAGTCACTCGGC 600
|
|
|
QY 707 CACTACCGGGGTCTCTTCCAGCACAGAGAGAGTGGCGCCCTCTTGTCTGGAG 766
|
|
|
Db 601 CACTACCGGGGTCTCTTCCAGCACAGAGAGAGTGGCGCCCTCTTGTCTGGAG 660
|
|
|
QY 767 GGGACCGGGTCCACACACACATCGGCCCTTAGAGAGCGAGCCATCTGGAG 822
|
|
|
Db 661 GGGACCGGGTCCACACACACATCGGCCCTTAGAGAGCGAGCCATCTGGAG 716
|
|
|
```

```
RESULT 3
BQ691705 655 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT 8046876 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6208939
5' mRNA sequence.
```

```
ACCESSION BQ691705
VERSION BQ691705.1 GI:21817021
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 655).
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2367 row: g column: 20
High quality sequence stop: 645.
Location/Qualifiers
1..655
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
FEATURES
source
```

```
/db_xref="taxon:9606"
/clone="IMAGE:6208939"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

```
BASE COUNT 155 a 216 c 193 g 91 t
ORIGIN
```

```
Query Match 61.5%; Score 652; DB 13; Length 655;
Best Local Similarity 100.0%; Pred. No. 4e-135;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 TACCTGCAGCAGCAGATCGACTGCGGCCACCATCTCGCTGTACAGCGGGAGGAGCCC 469
|
|
|
Db 1 TACCTGCAGCAGCAGATCGACTGCGGCCACCATCTCGCTGTACAGCGGGAGGAGCCC 60
|
|
|
QY 470 CCACCTTACAGGGGCCCTTGACCTTCAGCTTGGGACCCCGAGCAGCAGCTGGAATG 529
|
|
|
Db 61 CCACCTTACAGGGGCCCTTGACCTTCAGCTTGGGACCCCGAGCAGCAGCTGGAATG 120
|
|
|
QY 530 AACCGGAGTCTGTCGGCGCACCCCAACAGAACCATCTTCGACAGTCACTGTATGAT 589
|
|
|
Db 121 AACCGGAGTCTGTCGGCGCACCCCAACAGAACCATCTTCGACAGTCACTGTATGAT 180
|
|
|
QY 590 AGTGCCAGGCTGGGGCGGCCCTTGCCTCCAGCAGTAACTCGGGCATCAGCGCCACGTCG 649
|
|
|
Db 181 AGTGCCAGGCTGGGGCGGCCCTTGCCTCCAGCAGTAACTCGGGCATCAGCGCCACGTCG 240
|
|
|
QY 650 TACGCGACGCGCGGGCGCATGGAGGGCGCGCCCACTTACAGCGAGTCACTCGGCCAC 709
|
|
|
Db 241 TACGCGACGCGGGCGCATGGAGGGCGCGCCCACTTACAGCGAGTCACTCGGCCAC 300
|
|
|
QY 710 TACCGGGGTCTCTTCCAGCAGCAGCAGTGGCGGCCCTCTTGTCTGGAGGGG 769
|
|
|
Db 301 TACCGGGGTCTCTTCCAGCAGCAGCAGTGGCGGCCCTCTTGTCTGGAGGGG 360
|
|
|
QY 770 ACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGAGCAAGAG 829
|
|
|
Db 361 ACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGAGCAAGAG 420
|
|
|
QY 830 AAGGATAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGCGGCTGGGGCTGCG 889
|
|
|
Db 421 AAGGATAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGCGGCTGGGGCTGCG 480
|
|
|
QY 890 TAGTGAAAAGCGAGAACACTCGGGCTTCTTAGAAGAGGAGTGAGAGAGCGCGGGGG 949
|
|
|
Db 481 TAGTGAAAAGCGAGAACACTCGGGCTTCTTAGAAGAGGAGTGAGAGAGCGCGGGGG 540
|
|
|
QY 950 CGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTCTGTATATAATTTTACATGT 1009
|
|
|
Db 541 GCGAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTCTGTATATAATTTTACATGT 600
|
|
|
QY 1010 GATGCTGGTCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 1061
|
|
|
Db 601 GATGCTGGTCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 652
|
|
|
```

```
RESULT 4
AL558881/c
LOCUS AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS00J015YF12 3-PRIME, mRNA sequence.
ACCESSION AL558881
VERSION AL558881.2 GI:31283014
KEYWORDS EST.
```

```
951 bp mRNA linear EST 31-MAY-2003
```

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 951)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12903836.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 9945.r For more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DJ015DC06NP1&cluster=9945.r. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ InvitroGen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DJ015DC06NP1.  
 Location/Qualifiers  
 1..951  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DJ015YF12"  
 /cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
 /cell\_line="JURKAT"  
 /clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
 10-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and Bclor V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 141 a 275 c 318 g 201 t 16 others  
 ORIGIN  
 Query Match 61.5%; Score 652; DB 9; Length 951;  
 Best Local Similarity 99.6%; Pred. No. 2.9e-135;  
 Matches 802; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 168 TCACGTGCTGTGAGCCACTACAGTCTCTGCACGGTCTTCATCAGCGGCACAGCC 227  
 DB 846 TCACGTGCTGTGAGCCACTACAGTCTCTGCACGGTCTTCATCAGCGGCACAGCC 787  
 QY 228 AGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGAGCA 287  
 DB 786 AGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGAGCA 727  
 QY 288 CAGTGTACGGAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCTTGGCCCAACGACC 347  
 DB 726 CAGTGTACGGAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCTTGGCCCAACGACC 667  
 QY 348 GCTCGCTGCGCCCTTGGCCCGAGGAGCGTTCACCGCTTCAGAGCCACCTATC 407  
 DB 666 GCTCGCTGCGCCCTTGGCCCGAGGAGCGTTCACCGCTTCAGAGCCACCTATC 607  
 QY 408 CGTACCTGACGACGAGATCGACCTGCGCCGCCACCATCTTCGCTGTGACAGGGAGGAGC 467  
 DB 606 CGTACCTGACGACGAGATCGACCTGCGCCGCCACCATCTTCGCTTTACAGGGAGGAGC 547  
 QY 468 CCCACCTTACAGGGCCCTGACCTCCAGCTTCGGGACCCCGAGGAGAGCTTGGAAC 527  
 DB 546 CCCACCTTACAGGGCCCTGACCTCCAGCTTCGGGACCCCGAGGAGAGCTTGGAAC 487  
 QY 528 TGAACCGGGAGTGGTGGCGCACCCCAACAGAACCATCTTCACAGCTGACCTGATGG 587  
 DB 486 TGAACCGGGAGTGGTGGCGCACCCCAACAGAACCATCTTCACAGCTGACCTGATGG 427  
 QY 588 ATAGTGCCAGGTGGGCGGCCCTTGCCTCCCGCCAGCAGTAACTCGGGCATCAGCGCACGT 647  
 DB 426 ATAGTGCCAGGTGGGCGGCCCTTGCCTCCCGCCAGCAGTAACTCGGGCATCAGCGCACGT 367

QY 648 GCTAGCGCAGCGCGGGCGCATGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCC 707  
 DB 366 GCTAGCGCAGCGCGGGCGCATGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCC 307  
 QY 708 ACTACCGGGGTCTCTTCCAGCACCACGACGAGCAGTGGCGCCCTCTTGTGGAGG 767  
 DB 306 ACTACCGGGGTCTCTTCCAGCACCACGACGAGCAGTGGCGCCCTCTTGTGGAGG 247  
 QY 768 GGACCGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGGAGCAAG 827  
 DB 246 GGACCGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGGAGCAAG 187  
 QY 828 AGAAGGATAAAGAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGGCTG 887  
 DB 186 AGAAGGATAAAGAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGGCTG 127  
 QY 888 CGTAGGTGAAAGGAGCAACACTCCGCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGG 947  
 DB 126 CGTAGGTGAAAGGAGCAACACTCCGCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGG 67  
 QY 948 GCGCAGCAACGCGATCGTGTGGCCC 972  
 DB 66 GCGCAGCAACGCGATCGTGTGGCCC 42

RESULT 5  
 BQ636742 609 bp mRNA linear EST 15-JUL-2002  
 LOCUS hdl3h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he  
 DEFINITION Homo sapiens cDNA clone hdl3h06 5', mRNA sequence.  
 ACCESSION BQ636742  
 VERSION BQ636742.1 GI:21761201  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 609)  
 AUTHORS Wislow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman  
 J.W., Bouffard, G., Smith, D. and Peterson, K.  
 TITLE Expressed sequence tag analysis of human retina for the NEIBank  
 project: Retbindin, an abundant, novel retinal cDNA and alternative  
 splicing of other retina-preferred gene transcripts  
 JOURNAL Mol. Vis. 8 (4), (2002) In press  
 COMMENT Contact: Wislow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 13 row: h column: 06  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..609  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hdl3h06"  
 /tissue\_type="Retina"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human Retina cDNA (Un-normalized, unamplified  
 ): hd/he"  
 /note="Organ: Eye; Vector: pSPORT1; Neural retina tissue  
 was dissected from two 80 year old donors with no observed  
 eye disease. 100ug of total RNA was used for library  
 construction. A directionally cloned cDNA library in the  
 pSPORT1 vector (Life Technologies) was constructed at  
 Bioserve Biotechnology (Laurel MD) essentially following  
 the protocols of the SuperScript plasmid System full  
 details of which are contained in the manufacturer's



Instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGAGCGGCGCC(T)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC).

114 a 238 c 182 g 75 t

Query Match 57.4%; Score 609; DB 13; Length 609;  
Best Local Similarity 100.0%; Pred. No. 1.3e-125;  
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 214 CAGCGGACACAGCGGCGGAGAGAGAGATGCTCTCAGAGAGTGTCTGTG 273
Db 1 CAGCGGACACAGCGGCGGAGAGAGATGCTCTCAGAGAGTGTCTGTG 60
QY 274 GCCTTCGAGAGACAGTGTAGGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCC 333
Db 61 GCCTTCGAGAGACAGTGTAGGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCC 120
QY 334 TCGGCCACCGACCGCTGCGCTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTT 393
Db 121 TCGGCCACCGACCGCTGCGCTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTT 180
QY 394 CCAGCCACCTATCGTACTCAGCAGAGATCGACCTGCGCCGCCACCATCTCGCTGC 453
Db 181 CCAGCCACCTATCGTACTCAGCAGAGATCGACCTGCGCCGCCACCATCTCGCTGC 240
QY 454 AGACGGGAGGAGCGCCACCTACCGAGCGCCCTGCACCTCCAGCTTCGGGACCCGA 513
Db 241 AGACGGGAGGAGCGCCACCTACCGAGCGCCCTGCACCTCCAGCTTCGGGACCCGA 300
QY 514 GCAGCAGCTGGAACCTGAAACCGGAGTGTGTGCGCGCACCCCAACAGAACCATCTCGA 573
Db 301 GCAGCAGCTGGAACCTGAAACCGGAGTGTGTGCGCGCACCCCAACAGAACCATCTCGA 360
QY 574 CAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
Db 361 CAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 634 CATCAGCGCAGCTGTCTACGGCAGCGGCGGCGCATGAGGAGCGCGCCGCCACCTACAG 693
Db 421 CATCAGCGCAGCTGTCTACGGCAGCGGCGGCGCATGAGGAGCGCGCCGCCACCTACAG 480
QY 694 CGAGGTCTATCGGCCACTACCGCGGCTCTCTTCCAGCACACAGCAGAGTGGCGCGCC 753
Db 481 CGAGGTCTATCGGCCACTACCGCGGCTCTCTTCCAGCACACAGCAGAGTGGCGCGCC 540
QY 754 CTCCTTGTGAGGGAGACCGGCTCCACACACATCGCGCCCTTAGAGAGCGCAGC 813
Db 541 CTCCTTGTGAGGGAGACCGGCTCCACACACATCGCGCCCTTAGAGAGCGCAGC 600
QY 814 CATCTGGAG 822
Db 601 CATCTGGAG 609
```

RESULT 6  
B0690750  
Locus  
DEFINITION B0690750 890 bp mRNA linear EST 15-JUL-2002  
5', mRNA sequence.  
AGENCOURT\_8046394 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6209341

ACCESSION B0690750  
VERSION B0690750.1 GI:21816066

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 890)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

# COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2368 row: h column: 14  
High quality sequence stop: 627.

# FEATURES

source

Location/Qualifiers  
1..890

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6209341"

/tissue\_type="ductal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_110"

/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GSCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT 168 a 303 c 273 g 146 t

# ORIGIN

Query Match 56.6%; Score 601; DB 13; Length 890;  
Best Local Similarity 99.8%; Pred. No. 5.5e-124;  
Matches 651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 TCCTCCTTGGCTTCGGTGAAGCGCTTGGGGTTCAGTGGGCCATGATCCCGAGCTGC 60
Db 24 TCCTCCTTGGCTTCGGTGAAGCGCTTGGGGTTCAGTGGGCCATGATCCCGAGCTGC 83
QY 61 TGGAGAACTGAAGCGGAGCGGTCTCTCGCAAAACAGGCAATGGCGGAGCTGAGTTTGT 120
Db 84 TGGAGAACTGAAGCGGAGCGGTCTCTCGCAAAACAGGCAATGGCGGAGCTGAGTTTGT 143
QY 121 TCAGATCATCATCATCTGTTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 180
Db 144 TCAGATCATCATCATCTGTTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 203
QY 181 GAGCCACTACAGCTGTCTGACGCTTCTATAGCCGGCACACCCAGGCGGCGGAGGAG 240
Db 204 GAGCCACTACAGCTGTCTGACGCTTCTATAGCCGGCACACCCAGGCGGCGGAGGAG 263
QY 241 AGAAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGGCAA 300
Db 264 AGAAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGGCAA 323
QY 301 CGGAATCCAGAGCGGAGGCTCTAGCGCCCGCTCGGCCACCCAGCCGCTGGCCGTGCC 360
Db 324 CGGAATCCAGAGCGGAGGCTCTAGCGCCCGCTCGGCCACCCAGCCGCTGGCCGTGCC 383
QY 361 GCCCTTCCCGCAGCGGAGCGCTTCCAGCGCTTCCAGCCACCTATCGTACTCGTGCAGCA 420
Db 384 GCCCTTCCCGCAGCGGAGCGCTTCCAGCGCTTCCAGCCACCTATCGTACTCGTGCAGCA 443
QY 421 CGAGATCGACCTGCGCCCGCCACCATCTCGCTGTACAGCGGAGGAGGAGCCCGCCCTTACCA 480
Db 444 CGAGATCGACCTGCGCCCGCCACCATCTCGCTGTACAGCGGAGGAGGAGCCCGCCCTTACCA 503
QY 481 GGGCCCTTCGACCTTCCAGCTTCGGGAGCCCGCAGCAGCAGCTGGAACCTGAACCGGAGTGC 540
Db 504 GGGCCCTTCGACCTTCCAGCTTCGGGAGCCCGCAGCAGCAGCTGGAACCTGAACCGGAGTGC 563
QY 541 GGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGCCTGTAGTGGATAGTGCAGGCT 600
```





```

|||||
457 CCCTGCCCCCCAGCAGTAACCTCGGGCATACGGCCACGCTACTCGGAGCGGGCGC 398
QY
668 ATGAGGGGGCCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGTCCTCTTC 727
Db
397 ATGAGGGGGCCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGTCCTCTTC 338
QY
728 CAGCACCGCAGCAGAGTGGGGCCCTCTTGTGAGGGGACCCGCTCCACCACACA 787
Db
337 CAGCACCGCAGCAGAGTGGGGCCCTCTTGTGAGGGGACCCGCTCCACCACACA 278
QY
788 CACATCGGGCCCTTAGAGCGCCAGCCATCTCGAGCAAGAGAGATATAACAGAAAGGA 847
Db
277 CACATCGGGCCCTTAGAGCGCCAGCCATCTCGAGCAAGAGAGATATAACAGAAAGGA 218
QY
848 CACCTCTCTTAGGGTCCCGAGGGGGCGGGTGTGGGTGTGAGTGAAGAGCAGAAC 907
Db
217 CACCTCTCTTAGGGTCCCGAGGGGGCGGGTGTGGGTGTGAGTGAAGAGCAGAAC 158
QY
908 ACTCGCGCTCTTAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
Db
157 ACTCGCGCTCTTAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 98
QY
968 GGCCTCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGTCTGTCTGAATGC 1027
Db
97 GGCCTCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGTCTGTCTGAATGC 38
QY
1028 ACAAGCTAAGAGAGCTTGCAGAAAAA 1057
Db
37 ACAAGCTAAGAGAGCTTGCAGAAAAA 8

```

```

RESULT 10
BE855409/c
LOCUS
DEFINITION
7913f05.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3306369 3'
similar to TR:Q9UJD3 Q9UJD3 D07187.1 ; mRNA sequence.
ACCESSION BE855409
VERSION BE855409.1 GI:10367404
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/STGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence scop: 443.
Location/Qualifiers
1. .559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3306369"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a

```

## FEATURES

```

source
1. .559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3306369"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Brn23"
/notes="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a

```

```

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCGATATCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 82 a 173 c 182 g 122 t
ORIGIN

```

```

Query Match 52.7%; Score 559; DB 10; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.5e-114;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 488 TGCACCTCTCAGCTTCGGGACCCGAGCAGCAGCTGGAACCTGAACCCGGAGTCTGGTGGCG 547
Db 559 TGCACCTCTCAGCTTCGGGACCCGAGCAGCAGCTGGAACCTGAACCCGGAGTCTGGTGGCG 500
QY 548 GCACCCCAACAGAACCATCTTCGACAGTGCACCTGATGGATAGTGCAGCTGGGCGCG 607
Db 499 GCACCCCAACAGAACCATCTTCGACAGTGCACCTGATGGATAGTGCAGCTGGGCGCG 440
QY 608 CCCTGCCCCCCAGCAGTAACCTCGGGCATCAGCGCCAGCTGCTACGGCAGCGCGCGCGCG 667
Db 439 CCCTGCCCCCCAGCAGTAACCTCGGGCATCAGCGCCAGCTGCTACGGCAGCGCGCGCGCG 380
QY 668 ATGAGGGGGCGCGCGCCACCTACAGCAGAGTCACTCGGCCACTACCCGGGTCCTCTTC 727
Db 379 ATGAGGGGGCGCGCGCCACCTACAGCAGAGTCACTCGGCCACTACCCGGGTCCTCTTC 320
QY 728 CAGCACCAGCAGCAGTGGGGCGCCCTCTTGTGGAGGGGACCCGCTCCACCACACA 787
Db 319 CAGCACCAGCAGCAGTGGGGCGCCCTCTTGTGGAGGGGACCCGCTCCACCACACA 260
QY 788 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATATAACAGAAAGGA 847
Db 259 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATATAACAGAAAGGA 200
QY 848 CACCTCTCTAGGTTCGCCAGGGGGCGGGTGTGGGTGTGGTGAAGAGCAGAAC 907
Db 199 CACCTCTCTAGGTTCGCCAGGGGGCGGGTGTGGGTGTGGTGAAGAGCAGAAC 140
QY 908 ACTCGCGCTCTTAGAGAGAGTGAAGAGAGCGGGGGCGCGCAGCAACGCTCGTGT 967
Db 139 ACTCGCGCTCTTAGAGAGAGTGAAGAGAGCGGGGGCGCGCAGCAACGCTCGTGT 80
QY 968 GGCCTCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGC 1027
Db 79 GGCCTCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGC 20
QY 1028 ACAAGCTAAGAGAGCTTGC 1046
Db 19 ACAAGCTAAGAGAGCTTGC 1

```

```

RESULT 11
BM974296/c
LOCUS
DEFINITION BM974296 626 bp mRNA linear EST 20-FEB-2003
UI-CF-EC1-aca-k-23-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
UI-CF-EC1-aca-k-23-0-UI 3', mRNA sequence.
ACCESSION BM974296
VERSION BM974296.1 GI:19591887
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 626)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)

```

MEDLINE	97044477
PUBMED	889548
COMMENT	Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics ( <a href="http://www.resgen.com">www.resgen.com</a> ) or from Open Biosystems ( <a href="http://www.openbiosystems.com">www.openbiosystems.com</a> ). Seq primer: M13 FORWARD

FEATURES	source
Location/Qualifiers	
1. .626	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="UI-CF-EC1-aca-k-23-0-UI"	
/tissue_type="Lung"	
/dev_stages="Adult and Fetal"	
/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"	
/clone_lib="UI-CF-EC1"	
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTGCCTAC."	
TAG_LIB=UI-CF-EC1	
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-377 and 380-383	
TAG_SEQ=AAGTGCCTAC"	
85 a	187 c 201 g 151 t 2 others
BASE COUNT	
ORIGIN	

Query Match	52.4%;	Score 556;	DB 12;	Length 626;
Best Local Similarity	100.0%;	Pred. No. 6.5e-114;		
Matches 556;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	506	GACCCCGAGCAGCTGGAACTGAAACGGGAGTCGGTGCGCCACCCCCAAACAGAACCC	565	
DB	559	GACCCCGAGCAGCTGGAACTGAAACGGGAGTCGGTGCGCCACCCCCAAACAGAACCC	500	
QY	566	ATCTTCGACAGTGCACCTGATGATAGTCCAGAGCTGGCGGGCCCTGCCCCCAGCAGT	625	
DB	499	ATCTTCGACAGTGCACCTGATGATAGTCCAGAGCTGGCGGGCCCTGCCCCCAGCAGT	440	
QY	626	AACTCGGGCATCAGCGCCACACGTGCTACGCGCAGCGGGCGGCATGAGGGCGCGCGCC	685	
DB	439	AACTCGGGCATCAGGGCCACGTGTACGGCAGCGGGCGGCATGAGGGCGCGCGCC	380	
QY	686	ACCTACAGCAGGTCATCGGCCATACCCCGGGTCTCTTTCAGCACACAGCAGACGAGT	745	
DB	379	ACCTACAGCAGGTCATCGGCCATACCCCGGGTCTCTTTCAGCACACAGCAGACGAGT	320	
QY	746	GGGCGGCCCTCTTGCTGTGAGGGGACCGCGTCTCACACACACATCGCGCCCTTAGAG	805	
DB	319	GGGCGGCCCTCTTGCTGTGAGGGGACCGCGTCTCACACACACATCGCGCCCTTAGAG	260	

Qy	806	AGCGACCCATCTGGAGCAAAAGAGAGAGTAAACAGAAAGGACACCTCTCTTAGGGTCCC	865
Db	259	AGCGACCCATCTGGAGCAAAAGAGAGAGTAAACAGAAAGGACACCTCTCTTAGGGTCCC	200
Qy	866	CAGGGGGCGGGCTCGGGCTCGTAGGTGAAAAGGCAGAACTCCGCGCTTCTTAGAA	925
Db	199	CAGGGGGCGGGCTCGGGCTCGTAGGTGAAAAGGCAGAACTCCGCGCTTCTTAGAA	140
Qy	926	GAGGAGTGAGAGGAAGCGGGGGGCGCAGCAACGCATCTGTGTGGCCCTCCCTCCACCT	985
Db	139	GAGGAGTGAGAGGAAGCGGGGGGCGCAGCAACGCATCTGTGTGGCCCTCCCTCCACCT	80
Qy	986	CCCTGTGTTAAATATTACATGTGATCTGTCTGAATGCAAGCTTAAGAGAGCTTG	1045
Db	79	CCCTGTGTTAAATATTATTACATGTGATCTGTCTGAATGCAAGCTTAAGAGAGCTTG	20
Qy	1046	CAAAAAAAAAAAAAA	1061
Db	19	CAAAAAAAAAAAAAA	4

RESULT 12

BO575582/c

LOCUS

BO575582

UI-H-EZ1-bbf-f-09-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone

UI-H-EZ1-bbf-f-09-0-UI 3', mRNA sequence.

ACCESSION

BO575582

VERSION

BO575582.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 570)

TITLE

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of Orthopaedics  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLYA=Yes.

```

FEATURES             Location/Qualifiers
     1..570
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-H-EZ1-bbf-f-09-0-UI"
     /tissue_type="Chondrosarcoma Grade II"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies)"
     /clone_lib="NCI CGAP Ch2"
     /note="Organ: Left Pélvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI CGAP Ch2 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma Grade II. The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(AT)18 tail. The sequence tag for this library is

```

TGATCAGCT.  
TAG LIB=UI-H-E21  
TAG\_TISSUE=grade-2\_chondrosarcoma  
TAG\_SEQ=ATCTAATATG

BASE COUNT 80 a 174 c 177 g 139 t  
ORIGIN

Query Match 50.5%; Score 536; DB 13; Length 570;  
Best Local Similarity 100.0%; Pred. No. 1.8e-109;  
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 522 TGGAACTGAACCGGAGTCGGTGGCGGACACCCCAAGAACCACTCTTCGACAGTGACC 581  
Db 543 TGGAACTGAACCGGAGTCGGTGGCGGACACCCCAAGAACCACTCTTCGACAGTGACC 484  
QY 582 TGATGATAGTCGCCAGCTGGCGGCGCCCTGCCCCCAAGAGTAACTCGGGCATCAGCG 641  
Db 483 TGATGATAGTCGCCAGCTGGCGGCGCCCTGCCCCCAAGAGTAACTCGGGCATCAGCG 424  
QY 642 CCACGTGCTACGGCAGCGGCGGCGCATGAGAGGCGCGCCACCTACAGCGAGGTCA 701  
Db 423 CCACGTGCTACGGCAGCGGCGGCGCATGAGAGGCGCGCCACCTACAGCGAGGTCA 364  
QY 702 TCGGCCACTACCGGGGTCCTCTTCCAGCACAGCAGAGTGGCGGCCCTCTCTTGC 761  
Db 363 TCGGCCACTACCGGGGTCCTCTTCCAGCACAGCAGAGTGGCGGCCCTCTCTTGC 304  
QY 762 TGGAGGGACCGGGTCCACACACACATCGCGCCCTAGAGAGCGAGCATCTGA 821  
Db 303 TGGAGGGACCGGGTCCACACACACATCGCGCCCTAGAGAGCGAGCATCTGA 244  
QY 822 GCAGAGAGAGTAAACAGAGAGACACCTCTTAGGGTCCCGAGGGGCGGGCTG 881  
Db 243 GCAGAGAGAGTAAACAGAGAGACACCTCTTAGGGTCCCGAGGGGCGGGCTG 184  
QY 882 GGGCTGCGTAGTGAAGAGAGACACCTCCCGCTTCTTAGAGAGGAGTGAAGAGAG 941  
Db 183 GGGCTGCGTAGTGAAGAGAGACACCTCCGGCTTCTTAGAGAGGAGTGAAGAGAG 124  
QY 942 GCGGGGGCGAGCAACGATCGTGTGGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1001  
Db 123 GCGGGGGCGAGCAACGATCGTGTGGCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 64  
QY 1002 TTACATGATCTGTGCTGAATGCAAGTGAAGAGAGTTCGAAAAA 1057  
Db 63 TTACATGATCTGTGCTGAATGCAAGTGAAGAGAGTTCGAAAAA 8

RESULT 13  
BU539219  
LOCUS  
DEFINITION  
AGENCOURT\_10215265 NIH\_MGC\_107 Homo sapiens cDNA clone  
IMAGE:4569922 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 945)  
NIH-MGC <http://mgi.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LCM2757 row: p column: 18  
High quality sequence stop: 663.  
FEATURES  
Location/Qualifiers  
1..945  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4569922"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 194 a 319 c 287 g 144 t 1 others  
ORIGIN

Query Match 50.5%; Score 536; DB 13; Length 945;  
Best Local Similarity 99.5%; Pred. No. 1.2e-109;  
Matches 736; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 18 TGAAGCGCTTGGGGGTTCACTGGGCCATGATCCCCAGCTGCTGGAGAACTGAAGCGG 77  
Db 1 TGAAGCGCTTGGGGGTTCACTGGGCCATGATCCCCAGCTGCTGGAGAACTGAAGCGG 60  
QY 78 ACGGTCTCTCGAAACAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCG 137  
Db 61 ACGGTCTCTCGAAACAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCG 120  
QY 138 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 197  
Db 121 TGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 198 CTGACCGTCTCTTATAGCCGGCACAGCAGGGGCGGAGGAGAGAAATGCTTCTCT 257  
Db 181 CTGACCGTCTCTTATAGCCGGCACAGCAGGGGCGGAGGAGAAATGCTTCTCT 240  
QY 258 CAGAGGATGCTGTGGCCCTCGGAGAGACAGTGTAGGCAACGGAATCCAGAGCGCG 317  
Db 241 CAGAGGATGCTGTGGCCCTCGGAGAGACAGTGTAGGCAACGGAATCCAGAGCGCG 300  
QY 318 AGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 377  
Db 301 AGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 360  
QY 378 AGCGTTCCACCGCTTCCAGCCCACTATCGTACCTGACGACGAGATCGACCTGCGCG 437  
Db 361 AGCGTTCCACCGCTTCCAGCCCACTATCGTACCTGACGACGAGATCGACCTGCGCG 420  
QY 438 CCACCATCTCGTGTGACAGCGGAGGAGCCCCACCTTACCGAGGCGCTTGCACCTCC 497  
Db 421 CCACCATCTCGTGTGACAGCGGAGGAGCCCCACCTTACCGAGGCGCTTGCACCTCC 480  
QY 498 AGCTTCGGGACCCCGAGCAGCAGCTGGAATGAACCGGAGTTCGGTGGCGCACCCCCAA 557  
Db 481 AGCTTCGGGACCCCGAGCAGCAGCTGGAATGAACCGGAGTTCGGTGGCGCACCCCCAA 540  
QY 558 ACAGAACCATCTTCCAGAGTACCTGATGAGTGGCCAGCTGGCGGCGCTTGCCTCC 617  
Db 541 ACAGAACCATCTTCCAGAGTACCTGATGAGTGGCCAGCTGGCGGCGCTTGCCTCC 600  
QY 618 CCACAGTAACTCGGGCATCAGCGCATGCTGCTAGCGAGCGGCGGCGCATGGAGGGG 677  
Db 601 CCACAGTAACTCGGGCATCAGCGCATGCTGCTAGCGAGCGGCGGCGCATGGAGGGG 660  
QY 678 CGCGCCCACTTACAGCGAGGTCATCGGCCACTACCGGGGTCCTCTTTCAGACACAGC 737  
Db 661 CGCGCCCACTTACAGCGAGGTCATCGGCCACTACCGGGGTCCTCTTTCAGACACAGC 720

```

QY 738 AGACAGTGGGGCGCCCTCC 757
Db 721 AGAGCAGTGGGGCGCCCTCC 740

RESULT 14
LOCUS CB049800
DEFINITION NISC_G113c01.y1 NCI_CGAP_Pr28 Homo sapiens mRNA linear EST 17-JAN-2003,
5', mRNA sequence.
ACCESSION CB049800
VERSION CB049800.1 GI:27788087
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM8009 row: E column: 1
Seq primer: ML3RP1 reverse primer (ABI).
FEATURES
Location/Qualifiers
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3271656"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 146 a 170 c 168 g 79 t
ORIGIN
Query Match 50.0%; Score 530; DB 14; Length 563;
Best Local Similarity 100.0%; Pred. No. 3.9e-108;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 TGAACCGGAGTCGTGGCGGCACCCCAACAGAACCATCTTCACAGTGCACCTGATGG 587
Db 10 TGAACCGGAGTCGTGGCGGCACCCCAACAGAACCATCTTCACAGTGCACCTGATGG 69

QY 588 ATAGTGCAGGCTGGGGCCCTCCCGCCAGCAGTAACCTCGGGCATCAGCGCAGT 647
Db 70 ATAGTGCAGGCTGGGGCCCTCCCGCCAGCAGTAACCTCGGGCATCAGCGCAGT 129

QY 648 GCTACGCGAGCGGGCGCGATGGAGGGCGCCCGCCACCTACAGCGAGGTTCATCGGCC 707
Db 130 GCTACGCGAGCGGGCGCGATGGAGGGCGCCCGCCACCTACAGCGAGGTTCATCGGCC 189

QY 708 ACTACCCGGGGTCTCTCTTCCAGCACACAGCAGAGCAGTGGGGCCCTCTTGTGAGG 767

```

```

Db 190 ACTACCCGGGGTCTCTCTTCCAGCACACAGCAGTGGGGCCCTCTTGTGAGG 249
QY 768 GGACCCGGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGACAAAG 827
Db 250 GGACCCGGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGACAAAG 309
QY 828 AGAAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTG 887
Db 310 AGAAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTG 369
QY 888 CGTAGGTGAAGAGCAGAACACTCGCGCTTCTTAGAGAGGAGTGAGAGAGCGCGGG 947
Db 370 CGTAGGTGAAGAGCAGAACACTCGCGCTTCTTAGAGAGGAGTGAGAGAGCGCGGG 429
QY 948 GCGCAGCAACGCGCATCGTGGCGCCCTCCCTCCACCTCCCTGTGTATAAATATTACAT 1007
Db 430 GCGCAGCAACGCGCATCGTGGCGCCCTCCCTCCACCTCCCTGTGTATAAATATTACAT 489
QY 1008 GTGATGTCTGTCTGAATGCAAGACTTAAGAGAGCTTCAAAAAA 1057
Db 490 GTGATGTCTGTCTGAATGCAAGACTTAAGAGAGCTTCAAAAAA 539

RESULT 15
LOCUS BQ686793
DEFINITION AGENCOURT 8345390 NIH_MGC_110 Homo sapiens linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION BQ686793
VERSION BQ686793.1 GI:21812109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2387 row: e column: 05
High quality sequence stop: 593.
FEATURES
Location/Qualifiers
1..844
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6248044"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 239 a 236 c 235 g 133 t 1 others
ORIGIN
Query Match 49.8%; Score 528; DB 13; Length 844;
Best Local Similarity 99.8%; Pred. No. 7.6e-108;

```

Matches 578;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	483	GCCCTGACCCCTCAGCTTCGGGACCCGACAGCAGCTGGAACCTGAACCGGGAGTCGG 542							
Db	1								
		1 GCCCTGACCCCTCAGCTTCGGGACCCGACAGCAGCTGGAACCTGAACCGGGAGTCGG 60							
Qy	543	TGCGGCACCCCAACAGAACATCTTCGACAGTGACCTGATGATAGTGCAGGCTGG 602							
Db	61								
		TGCGGCACCCCAACAGAACATCTTCGACAGTGACCTGATGATAGTGCAGGCTGG 120							
Qy	603	GGGGCCCTCGCCGCCACAGTAACCTCGGGCATCAGGCCACGTGCTACGGCAGCGCG 662							
Db	121								
		GGGGCCCTCGCCGCCACAGTAACCTCGGGCATCAGGCCACGTGCTACGGCAGCGCG 180							
Qy	663	GGGGCATGGAGGGCGCGCCGACCTACAGCAGGTGATCGGCCACTACCGGGGTCTT 722							
Db	181								
		GGGGCATGGAGGGCGCGCCGACCTACAGCAGGTGATCGGCCACTACCGGGGTCTT 240							
Qy	723	CCTTCAGCACAGCAGCAGTGGCGCCCTCTTGTGGAGGGGACCGGGTCCACC 782							
Db	241								
		CCTTCAGCACAGCAGCAGTGGCGCCCTCTTGTGGAGGGGACCGGGTCCACC 300							
Qy	783	ACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGA 842							
Db	301								
		ACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGA 360							
Qy	843	AAGGACACCTCTCTAGGTCCCGAGGGGGCGGGCTGGGGCTGGCTAGGTGAAAAGGC 902							
Db	361								
		AAGGACACCTCTCTAGGTCCCGAGGGGGCGGGCTGGGGCTGGCTAGGTGAAAAGGC 420							
Qy	903	AGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGAGGGCGGGGGCGCAGCAACGCAT 962							
Db	421								
		AGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGAGGGCGGGGGCGCAGCAACGCAT 480							
Qy	963	CGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATTTACATGTGATGTCTGTCTG 1022							
Db	481								
		CGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATTTACATGTGATGTCTGTCTG 540							
Qy	1023	AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAAA 1061							
Db	541								
		AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAAA 579							

Search completed: December 8, 2003, 06:17:44  
Job time : 2648 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 21:17:14 ; Search time 344 Seconds  
(without alignments)  
8325.883 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctccctgggttcgggtga.....cttgcaaaaaaaaaa 1061

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq 13Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1061	100.0	1061	21	Sequence encoding
2	1057.8	99.7	1334	25	Human secretory po
3	1034.4	97.5	1140	24	Prostate cancer-as
4	1021.2	96.2	1066	22	Human polynucleoti
5	991.2	93.4	1069	22	Human polynucleoti
6	961.8	90.7	969	21	cDNA encoding a hu
7	960.2	90.5	969	21	cDNA clone encodin
8	960.2	90.5	969	21	cDNA clone encodin

9	960.2	90.5	969	21	AAA75165	cDNA clone encodin
10	800.6	75.5	1321	24	ABK12137	Human cDNA encodin
11	728.8	68.7	1583	24	ABSG1424	Prostate specific
12	644.2	60.7	693	24	ABK12143	Human MIVR-1 homol
13	523.8	49.4	878	24	ABK12142	Mouse cDNA encodin
14	521.4	49.1	1713	21	AAA75152	cDNA encoding a mu
15	519.8	49.0	1713	21	AAA75166	cDNA clone encodin
16	519.8	49.0	1713	21	AAA75167	cDNA clone encodin
17	519.8	49.0	1713	21	AAA75168	cDNA clone encodin
18	477.8	45.0	1879	23	AAAS4503	DNA encoding novel
19	418.6	39.5	474	25	ABZ84732	Toxicologically re
20	399.8	37.7	408	22	AAFG5983	Novel human polynu
21	399	37.6	673	24	ABT09178	Phase-1 Rat CT gen
22	365.6	34.5	812	20	AAZ52964	Human prostate tum
23	352.4	33.2	937	21	AAZ52507	Human secreted pro
24	350.8	33.1	920	24	ABK34251	Human cDNA for nov
25	350	33.0	8093	24	ABK12145	Human MIVR-1 homol
26	349	32.9	8494	23	AAV77304	DNA encoding novel
27	345.8	32.6	8065	19	AAV38335	Manic-depressive i
28	334.8	31.6	8103	23	AAV77312	DNA encoding novel
29	281	26.5	8440	23	AAV77305	DNA encoding novel
30	281	26.5	8440	24	ABK83477	Human cDNA differe
31	277.8	26.2	8011	19	AAV38336	Manic-depressive i
32	253.6	23.9	8041	23	AAV77310	DNA encoding novel
33	249.2	23.5	475	24	ABK12144	Human MIVR-1 homol
34	239.4	22.6	426	23	AAAS4502	DNA encoding novel
35	232	21.9	254	21	AAA41265	Human secreted exp
36	223.4	21.1	1428	23	AAAS92497	DNA encoding novel
37	223.4	21.1	2942	23	AAV77313	DNA encoding novel
38	188.4	17.8	522	24	ABT10027	Human breast cance
39	60	5.7	60	24	ABN40872	Human spliced tran
40	52.8	5.0	114955	20	AAAS3491	Human adenosine Al
41	51	4.8	51	22	AAH89714	Human coding sequ
42	48.4	4.6	1337	20	AAZ17263	Human gene express
43	48.4	4.6	11304	22	AAI03113	Human reproductive
44	48.2	4.5	336	24	ABSA4598	DNA encoding varia
45	48	4.5	114955	20	AAAS3491	Human adenosine Al

#### ALIGNMENTS

RESULT 1  
AAA47429  
ID AAA47429 standard; DNA; 1061 BP.  
XX  
AC AAA47429;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
Sequence encoding human neuron-associated protein.

Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
Alzheimer's disease; Pick's disease; Huntington's disease;  
dementia; Parkinson's disease; demyelinating disease; meningitis;  
prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
cerebral palsy; muscular dystrophy; central nervous system; CNS;  
peripheral nervous system; PNS; myopathy; schizophrenia;  
actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;  
cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
AIDS; Addison's disease; adult respiratory distress syndrome;  
allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
Werner syndrome; trauma; human; ds.

Hom sapiens.  
Location/Qualifiers  
101..859  
/\*tag= a  
/product= Neuron associated protein



KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
KW anti-inflammatory; immunosuppressive; neuroprotective; neurotropic;  
KW neuroleptic; anticonvulsant; cytoskeletal; antiparkinsonian; anxiolytic;  
KW antipsychotic; antianemic; anti-HIV; human immunodeficiency virus;  
KW secretory polynucleotide; secretory protein; gene; ss.

XX Homo sapiens.

OS WO200283876-A2.

PN 24-OCT-2002.

PD 27-MAR-2002; 2002WO-US09921.

PF 29-MAR-2001; 2001US-280067P.

PR 29-MAR-2001; 2001US-280068P.

PR 16-MAY-2001; 2001US-291280P.

PR 17-MAY-2001; 2001US-291829P.

PR 17-MAY-2001; 2001US-291849P.

PR 19-JUN-2001; 2001US-299428P.

PR 20-JUN-2001; 2001US-299776P.

PR 20-JUN-2001; 2001US-300001P.

XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amesley SR;

PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;

PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX WPI; 2003-075543/07.

DR P-PSDB; ABP75660.

XX New human secretory proteins and polynucleotides, useful for

PT diagnosing, treating or preventing autoimmune/inflammatory disorders

PT (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell

PT proliferations or cancers -

XX Claim 1; SEQ ID NO 267; 459pp + Sequence Listing; English.

XX The invention relates to a secretory polynucleotide (designated sptm)

CC comprising any of 567 polynucleotide sequences (AB235837-AB236403), a

CC naturally occurring polynucleotide sequence at least 90 % identical to

CC the polynucleotide sequence, a polynucleotide complementary to them or an

CC RNA equivalent of them. The polypeptide or polynucleotide are useful for

CC treating, preventing or diagnosing a disease or condition associated with

CC the expression of functional SPTM. These are particularly useful for

CC diagnosing, treating or preventing autoimmune/inflammatory disorders

CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,

CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,

CC schizophrenia or amnesia), or cell proliferative disorders (e.g.

CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,

CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,

CC breast, cervix or prostate).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1334 BP; 299 A; 381 C; 394 G; 260 T; 0 other;

XX Query Match 99.7%; Score 1057.8; DB 25; Length 1334;

XX Best Local Similarity 99.8%; Pred. No. 2.3e-219;

XX Matches 1059; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX 1 TCCTCTGGGTTCGGGTGAAGCGCTTGGGGTTTCAGTGGGCGCATGATCCCGAGCTGC 60

DB 1 TCCTCTGGGTTCGGGTGAAGCGCTTGGGGTTTCAGTGGGCGCATGATCCCGAGCTGC 60

QY 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAAACACAGCAATGGCGGAGCTGGAGTTGT 120  
DB 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAAACACAGCAATGGCGGAGCTGGAGTTGT 120  
QY 121 TCAGATCATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
DB 121 TCAGATCATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180  
QY 181 GAGCCACTACAAGTGTCTGACCGGTCTTTCATCAGCCGGACACAGGCGGAGGAG 240  
DB 181 GAGCCACTACAAGTGTCTGACCGGTCTTTCATCAGCCGGACACAGGCGGAGGAG 240  
QY 241 AGAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCGAGAGACACAGTGTGAGGCAA 300  
DB 241 AGAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCGAGAGACACAGTGTGAGGCAA 300  
QY 301 CGGATATCCAGAGCGGACGGTCTAGCGCCCGCTCGGCGCCACCGACCGCTGGCGTGC 360  
DB 301 CGGATATCCAGAGCGGACGGTCTAGCGCCCGCTCGGCGCCACCGACCGCTGGCGTGC 360  
QY 361 GCGCTTCCCGCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCGTACCTCAGCA 420  
DB 361 GCGCTTCCCGCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCGTACCTCAGCA 420  
QY 421 CGAGATGACCTGCGCCCGCTTCCACCGCTTCCAGCCACCTATCGTACCTCAGCA 480  
DB 421 CGAGATGACCTGCGCCCGCTTCCACCGCTTCCAGCCACCTATCGTACCTCAGCA 480  
QY 481 GGGCCCTGTCACCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGAACTGAAACCGGAGTC 540  
DB 481 GGGCCCTGTCACCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGAACTGAAACCGGAGTC 540  
QY 541 GGTGGCGCACCCCGCAACAGAACCATCTTCGACAGTGCACCTGTAGTGGTGGTGGTGGT 600  
DB 541 GGTGGCGCACCCCGCAACAGAACCATCTTCGACAGTGCACCTGTAGTGGTGGTGGTGGT 600  
QY 601 GGGGGGCGCTTGGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGAGCGG 660  
DB 601 GGGGGGCGCTTGGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGAGCGG 660  
QY 661 GGGGGGCGATGAGGGGCGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGG 720  
DB 661 GGGGGGCGATGAGGGGCGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGG 720  
QY 721 CTCCTTCAGACACAGCAGCAGCAGTGGGCGCGCTTCTTGTGGAGGGAGCCCGGCTCCA 780  
DB 721 CTCCTTCAGACACAGCAGCAGCAGTGGGCGCGCTTCTTGTGGAGGGAGCCCGGCTCCA 780  
QY 781 CCACACACATCGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGAGGATAAACA 840  
DB 781 CCACACACATCGCGCCCGCTTCCAGCTTGGGACCCCGAGCAGCAGCTGGTACGGAGGATAAACA 840  
QY 841 GAAAGGACACCTCTCTAGAGTCCCGAGGGGCGCGGCTGGGGCTGGTGGTGGTGGTGGTGGT 900  
DB 841 GAAAGGACACCTCTCTAGAGTCCCGAGGGGCGCGGCTGGGGCTGGTGGTGGTGGTGGTGGT 900  
QY 901 GCAGAACTCCGGCTTCTTGAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
DB 901 GCAGAACTCCGGCTTCTTGAAGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
QY 961 ATCGTGTGGCGCTTCCCGCTTCCAGCTTGGTGTGTAATAATTTTACATGTGTGTGTGTGTGT 1020  
DB 961 ATCGTGTGGCGCTTCCCGCTTCCAGCTTGGTGTGTAATAATTTTACATGTGTGTGTGTGTGT 1020  
QY 1021 TGAATGCAAGCTAAGAGAGCTTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1061  
DB 1021 TGAATGCAAGCTAAGAGAGCTTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1061

RESULT 3

ABK92120

ID ABK92120 standard; DNA; 1140 BP.

XX AC ABK921120;  
 XX DT 15-AUG-2002 (first entry)  
 XX DE Prostate cancer-associated DNA sequence #6.  
 XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;  
 XX KW gene therapy; gene; ds.  
 XX OS Mammalia.  
 XX PN WO200230268-A2.  
 XX PD 18-APR-2002.  
 XX PF 12-OCT-2001; 2001WO-US32045.  
 XX PR 13-OCT-2000; 2000US-0687576.  
 XX PR 08-DEC-2000; 2000US-0733288.  
 XX PR 08-DEC-2000; 2000US-0733742.  
 XX PR 24-JAN-2001; 2001US-263957P.  
 XX PR 16-MAR-2001; 2001US-276791P.  
 XX PR 16-MAR-2001; 2001US-276888P.  
 XX PR 06-APR-2001; 2001US-281922P.  
 XX PR 24-APR-2001; 2001US-286214P.  
 XX PR 30-APR-2001; 2001US-0847046.  
 XX PR 04-MAY-2001; 2001US-288589P.  
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 XX DR P-PSDB; ABG61805.  
 XX DR WPI; 2002-471335/50.  
 XX DR P-PSDB; ABG61805.  
 XX PT Detecting a prostate cancer-associated transcript in a cell in a  
 PT patient, useful for diagnosing prostate cancer (PC) or screening  
 PT modulators of PC, by determining if prostate cancer-associated genes  
 PT are expressed in a prostate tissue.  
 XX PS Claim 22; Page 305; 436pp; English.  
 XX CC The present invention relates to methods of detecting a prostate  
 CC cancer-associated transcript in a cell from a patient. The method  
 CC comprises contacting a biological sample from the patient with  
 CC prostate cancer-associated polynucleotides (designated PC genes) that  
 CC selectively hybridise to a sequence that is at least 80% identical  
 CC to them. The prostate cancer-associated polynucleotide sequences  
 CC are differentially expressed in prostate tumour tissue or in  
 CC prostate cancer and are derived from the tissues of various  
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).  
 CC The methods of the invention are useful for diagnosing and treating  
 CC prostate cancer in mammals. The prostate cancer-associated genes are  
 CC useful for diagnosing or treating prostate cancer, as well as for  
 CC identifying modulators of prostate cancer or agents that inhibit  
 CC prostate cancer. The nucleic acid sequences are particularly useful  
 CC in gene therapy, as a vaccine or in antisense applications.  
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide  
 CC sequences.  
 XX SQ Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;  
 Query Match 97.5%; Score 1034.4; DB 24; Length 1140;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-214;  
 Matches 1050; Conservative 0; Mismatches 6; Indels 2; Gaps 1;  
 QY 4 TCCTTGGGTTCCGGTGAAGCCCTTGGGGTTTCAGTGGCCCATGATCCCCGAGCTGCTGG 63  
 DB 1 TCCTTGGGTTCCGGTGAAGCCCTTGGGGTTTC--GTGGCCCATGATCCCCGAGCTGCTGG 58  
 QY 64 AGAAGTGAAGGGGGGGTCTCTCTGCGAAACAGGCAATGGCGGAGCTGGAGCTTTGTTCA 123

59 AGAAGTGAAGGGGGAGACAGTCTCTCTGCGAAACAGGCAATGGCGGAGCTGGAGTTGTTCA 118  
 124 GATCATCATATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 183  
 119 GATCATCATATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 178  
 184 CCATCTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACAGCCAGGCGGCGAGGAGAGA 243  
 179 CCATCTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACAGCCAGGCGGCGAGGAGAGA 238  
 244 AGATGCCCTCTCTCAGAAAGGATGCTCTGTGGCCCTTCGAGAGACACAGTGTTCAGGCAACGG 303  
 239 AGATGCCCTCTCTCAGAAAGGATGCTCTGTGGCCCTTCGAGAGACACAGTGTTCAGGCAACGG 298  
 304 AATCCAGAGCGGAGGCTTACGCCCTTCGGCCCAACCGACCGCTGGCGGCTGGCGGCC 363  
 299 AATCCAGAGCGGAGGCTTACGCCCTTCGGCCCAACCGACCGCTGGCGGCTGGCGGCC 358  
 364 CTTCGCCAGCGGAGGCGCTTCCACCGCTTCCAGCCACCTTATCGTACCTGCAGACAGA 423  
 359 CTTCGCCAGCGGAGGCGCTTCCACCGCTTCCAGCCACCTTATCGTACCTGCAGACAGA 418  
 424 GATCGACCTGCGCCCAACCATCTCGTGTTCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 483  
 419 GATCGACCTGCGCCCAACCATCTCGTGTTCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 478  
 484 CCCCTGACCTTCAGCTTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTGGT 543  
 479 CCCCTGACCTTCAGCTTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTGGT 538  
 544 GCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATG 603  
 539 GCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATG 598  
 604 CGGCCCTGCGCCCAACAGAACCATCTTCGAGCGCATCAGCGCCACCTGCTACGCGAGCGGG 663  
 599 CGGCCCTGCGCCCAACAGAACCATCTTCGAGCGCATCAGCGCCACCTGCTACGCGAGCGGG 658  
 664 CGCATGAGGAGGCGCGCCCAACCTTACGCGAGTATTCGGCCACTACCGGGGGTCTCTC 723  
 659 CGCATGAGGAGGCGCGCCCAACCTTACGCGAGTATTCGGCCACTACCGGGGGTCTCTC 718  
 724 CTTCAGCACAGCAGCAGCAGTGGGCGCCCTCTTCTGCTGGAGGGGACCCCGCTCCACCA 783  
 719 CTTCAGCACAGCAGCAGCAGTGGGCGCCCTCTTCTGCTGGAGGGAGCCCGCTCCACCA 778  
 784 CACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAG 843  
 779 CACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAG 838  
 844 AGGACACCTCTTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGGTAGGTGAAAAGGCA 903  
 839 AGGACACCTCTTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGGTAGGTGAAAAGGCA 898  
 904 GAACACTCCGCGCTTCTTAGAGAGAGAGTGAAGAGCGGGGGCGCGAGCAACGCATC 963  
 899 GAACACTCCGCGCTTCTTAGAGAGAGAGTGAAGAGCGGGGGCGCGAGCAACGCATC 958  
 964 GTGTGGCCCTCCCTCCACCTCCCTGCTGTATAAATATTTACATGTGATGTGCTCTGA 1023  
 959 GTGTGGCCCTCCCTCCACCTCCCTGCTGTATAAATATTTACATGTGATGTGCTCTGA 1018  
 1024 ATGCACAGCTAAGAGAGCTTGCAAAAAAAAAAAAAAAAAAAAA 1061  
 1019 ATGCACAGCTAAGAGAGCTTGCAAAAAAAAAAAAAAAAAAAAA 1056  
 RESULT 4  
 AAI57868  
 ID AAI57868 standard; cDNA; 1066 BP.  
 XX AC  
 XX AC AAI57868;  
 XX XX









AC AAA75163;  
 XX 15-JAN-2001 (first entry)  
 XX cDNA clone encoding a human TANGO 261 polypeptide.  
 DE TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;  
 KW cellular proliferation; cellular differentiation; cellular adhesion;  
 KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
 KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
 KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
 KW intestinal disorder; spleen associated disease; renal disorder;  
 KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
 KW brain herniation; iatrogenic disease; inflammation; meningitis;  
 KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
 KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX 6..784  
 XX CDS /\*tag= a  
 XX /product= "TANGO 261"  
 XX WO200052022-A1.  
 XX  
 XX 08-SEP-2000.  
 XX  
 XX 01-MAR-2000; 2000WO-US05226.  
 XX  
 XX 01-MAR-1999; 99US-0122458.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
 XX  
 XX WPI; 2000-579269/54.  
 XX P-PSDB; AAB18461.  
 XX  
 XX Novel human and murine secreted proteins designated TANGO 216, 261,  
 XX 262, 266 and 267 useful as modulating agents of cellular processes,  
 XX e.g. for treating cancer -  
 XX  
 XX Disclosure; Page -; 175pp; English.  
 XX  
 XX AAA75163-65 encode human TANGO 261 proteins. The specification also  
 XX describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO  
 XX polypeptides can be used to modulate cellular proliferation, modulate  
 XX cellular differentiation and/or modulate cellular adhesion. The  
 XX proteins can be used to treat any von Willebrand factor-associated  
 XX disorder, regulate extracellular matrix structuring, cellular adhesion,  
 XX and cell trafficking and/or migration, modulate cellular interactions,  
 XX modulate cell adhesion in proliferative disorders, such as cancer,  
 XX modulate the proliferation, differentiation, and/or function of cells  
 XX that appear in the bone marrow, and leukocytes, treat bone marrow, blood  
 XX and hematopoietic associated diseases and disorders, atelectasis, bronchial  
 XX pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial  
 XX asthma and bronchiectasis, intestinal disorders, spleen associated  
 XX diseases, modulate renal disorders, treat cardiovascular disorders such  
 XX as ischemic heart disease, modulate the proliferation, differentiation,  
 XX and/or function of bone and cartilage cells and to treat bone and/or  
 XX cartilage associated diseases or disorder. They may also be used to  
 XX treat disorders associated with the ovaries, and cerebral oedema,  
 XX hydrocephalus, brain herniations, iatrogenic disease, inflammations,  
 XX bacterial and viral meningitis, Alzheimer's Disease, cerebral  
 XX toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,  
 XX hydrocephalus and encephalitis, and treat hepatic disorders.  
 XX note: the present sequence does not appear in the specification; it was  
 XX created using information provided.

XX Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 90.5%; Score 960.2; DB 21; Length 969;

Best Local Similarity 99.7%; Pred. No. 2.8e-198;  
 Matches 962; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 97 GGCATATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGTGATGATGATGAT 156  
 DB 2 GGAGATGGCGGACCTGGAGTTTGTTCAGATCATCATCATCGTGGTGTGATGATGATGAT 61  
 QY 157 GGTGTGTGTATCAGTCCCTGCTGAGCCACTACAGCTGTCTGACCGTCTTCATCAG 216  
 DB 62 GGTGTGTGTATCAGTCCCTGCTGAGCCACTACAGCTGTCTGACCGTCTTCATCAG 121  
 QY 217 CCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCTCAGAAAGATGCTGTGGCC 276  
 DB 122 CCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCTCAGAAAGATGCTGTGGCC 181  
 QY 277 CTCGGAGAGCAGTGTTCAGGCAACCGAATCCAGAGCGGAGGTCTACGCCGCCCTCG 336  
 DB 182 CTCGGAGAGCAGTGTTCAGGCAACCGAATCCAGAGCGGAGGTCTACGCCGCCCTCG 241  
 QY 337 GCCCAGCCAGCCCTGGCGGCTTGGCCCTTGGCCAGGGGAGCGCTTCCACGGCTTCCA 396  
 DB 242 GCCCAGCCAGCCCTGGCGGCTTGGCCCTTGGCCAGGGGAGCGCTTCCACGGCTTCCA 301  
 QY 397 GCCCAGCCATTCGTACCTGTCAGCAGAGATCGACCTGCGGCCACCATCTCTCGCTGT 456  
 DB 302 GCCCAGCCATTCGTACCTGTCAGCAGAGATCGACCTGCGGCCACCATCTCTCGCTGT 361  
 QY 457 CGGGAGAGAGCCGCCACCTTACCAGGGCCCTGTGACCTTCCAGTTCGGGAGCCCGAGCA 516  
 DB 362 CGGGAGAGAGCCGCCACCTTACCAGGGCCCTGTGACCTTCCAGTTCGGGAGCCCGAGCA 421  
 QY 517 CGAGCTGGAACCTGAAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTGCACAG 576  
 DB 422 CGAGCTGGAACCTGAAACCGGGAGTGGTGGCGGACCCCAACAGAACATCTTGCACAG 481  
 QY 577 TGACCTGATGATAGTGTGCGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGG 636  
 DB 482 TGACCTGATGATAGTGTGCGGCGGCTTGGCGGCTTGGCGGCTTGGCGGCTTGGCGG 541  
 QY 637 CAGGCCACGTGCTACCGCAGCGGCGGCGATGAGGGGCGCGCGCCACCTACAGCGA 696  
 DB 542 CAGGCCACGTGCTACCGCAGCGGCGGCGATGAGGGGCGCGCGCGCCACCTACAGCGA 601  
 QY 697 GGTCTATCGGCCACTACCGGGGCTCTCTCTCCAGCACAGAGAGAGTGGGCGGCCCTC 756  
 DB 602 GGTCTATCGGCCACTACCGGGGCTCTCTCTCCAGCACAGAGAGAGTGGGCGGCCCTC 661  
 QY 757 CTTGCTGAGGGGACCCCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGCAGCCAT 816  
 DB 662 CTTGCTGAGGGGACCCCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGCAGCCAT 721  
 QY 817 CTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTTAGGTCTCCAGGGGGCGG 876  
 DB 722 CTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTTAGGTCTCCAGGGGGCGG 781  
 QY 877 GCCTGGGCTGCTAGGTGAAAGGACACCTCCGCGCTTCTTAGAGAGAGTGA 936  
 DB 782 GCCTGGGCTGCTAGGTGAAAGGACACCTCCGCGCTTCTTAGAGAGAGTGA 841  
 QY 937 GGAAGCGGGGGCGGCGCAGCAACCGATCGTGTGGGCCCTCCCTCCACCTCTCTGTAT 996  
 DB 842 GGAAGCGGGGGCGGCGCAGCAACCGATCGTGTGGGCCCTCCCTCCACCTCTCTGTAT 901  
 QY 997 AATATTATCATGTGATGTCTGTGTAATGCAAGAGCTTAAGAGAGCTTGCATAAAAAA 1056  
 DB 902 AATATTATCATGTGATGTCTGTGTAATGCAAGAGCTTGCATAAAAAA 961  
 QY 1057 AAAAA 1061  
 DB 962 AAAAA 966

RESULT 8









XX WO200242776-A2.  
 XX 30-MAY-2002.  
 XX 01-NOV-2001; 2001WO-US45654.  
 XX 01-NOV-2000; 2000US-244782P.  
 XX (DIAD-) DIADEXUS INC.  
 XX Sun Y, Recipon H, Chen S, Liu C;  
 XX WPI; 2002-490217/52.  
 XX New polypeptide useful for diagnosing and monitoring the presence and  
 PT metastases of prostate cancer in a patient and as a component in  
 PT databases for search analysis as well as in sequence analysis  
 PT algorithms -  
 XX  
 PS Claim 1; Page 166-167; 242pp; English.  
 XX  
 CC The invention relates to an isolated polypeptide comprising a sequence  
 CC with 60 % identity to one of prostate specific protein (PSP) sequences,  
 CC or comprising an amino acid sequence encoded by one of 136 nucleotide  
 CC prostate specific nucleic acids sequences, PSNA, (or a sequence that  
 CC hybridises to it or is 60% identical to it), given in the specification.  
 CC Also included are a vector comprising the polynucleotide, a host cell  
 CC comprising the vector, an antibody specific for the PSP proteins and a  
 CC vaccine comprising the protein or polynucleotide. The PSP and PSNA are  
 CC useful for diagnosing and monitoring the presence and metastases of  
 CC prostate cancer in a patient. The PSNA is useful for determining the  
 CC the level PSNA in a sample. An antibody to the PSP is useful for  
 CC determining the presence of prostate specific protein in a sample, and  
 CC for treating a patient with prostate cancer, which induces an immune  
 CC response against the prostate cancer cell expressing the nucleic acid or  
 CC polypeptide and a kit is useful for detecting a risk of cancer or  
 CC presence of cancer in a patient. PSNA is useful as hybridisation probes  
 CC to detect, characterise and quantify hybridising nucleic acids from both  
 CC genomic and transcript-derived nucleic acid samples and also in  
 CC microarrays. Sequences of PSP and PSNA are useful as components in  
 CC databases for search analysis as well as in sequence analysis algorithms.  
 CC PSNA is useful to drive in vivo expression of PSP. The present  
 CC sequence is a PSNA of the invention.  
 XX  
 SQ Sequence 1583 BP; 361 A; 469 C; 457 G; 236 T; 0 other;  
 Query Match 68.7%; Score 728.8; DB 24; Length 1583;  
 Best Local Similarity 98.9%; Pred. No. 3.4e-148;  
 Matches 744; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 311 GAGCGCGAGGTCTAGCCCGCGCTCGGCCACCGACCGCTCGCGCTCGCGCTTCGCGC 370  
 DB |||||  
 825 GCGCGCGAGGTCTAGCCCGCGCTCGGCCACCGACCGCTCGCGCTTCGCGC 884  
 QY 371 CAGCGGAGCGCTTCACCGCTTCAGCCACCTACCTACCTACCGACGAGATCGAC 430  
 DB |||||  
 885 CAGCGGAGCGCTTCACCGCTTCAGCCACCTACCTACCTACCGACGAGATCGAC 944  
 QY 431 CTGCGCGCCACCATCTCGCTGTTCAGACGGGAGGAGCCCGCCCTACCGGCGCCCTGC 490  
 DB |||||  
 945 CTGCGCCACCATCTCGCTGTTCAGACGGGAGGAGCCCGCCCTACCGGCGCCCTGC 1004  
 QY 491 ACCCTCGAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGAGTCTGGTCGCGA 550  
 DB |||||  
 1005 ACCCTCGAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGAGTCTGGTCGCGA 1064  
 QY 551 CCCCAACACAGAACCATCTTCACAGTGACCTGATGATGATGATGATGATGATGATGATGAT 610  
 DB |||||  
 1065 CCCCAACACAGAACCATCTTCACAGTGACCTGATGATGATGATGATGATGATGATGATGAT 1124  
 QY 611 TGCCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGCGGCATG 670  
 |||||

Db 1125 TGCCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGCGAGCGGGCGGCATG 1184  
 QY 671 GAGGGCGCGCGCCACCTACAGCAGGTCTATCGCCCACTACCGGGGTCTCTCTTCCAG 730  
 Db 1185 GAGGGCGCGCGCCACCTACAGCAGGTCTATCGCCCACTACCGGGGTCTCTCTTCCAG 1244  
 QY 731 CACCAGCAGCAGTGGCGCCCTCTCTTGTGGAGGGGACCCGCTCCACACACACAC 790  
 Db 1245 CACCAGCAGCAGTGGCGCCCTCTCTTGTGGAGGGGACCCGCTCCACACACACAC 1304  
 QY 791 ATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGATAAAGAGAGAGAC 850  
 Db 1305 ATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGATAAAGAGAGAGAC 1364  
 QY 851 CCTCTTAGGCTCCCGAGGGGCGCGCTGGGCTGGGTAGGTGAAAGCAGACACT 910  
 Db 1365 CCTCTTAGGCTCCCGAGGGGCGCGCTGGGCTGGGTAGGTGAAAGCAGACACT 1424  
 QY 911 CCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGGCGACCAATCGCATGCGTGG 969  
 Db 1425 CCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGGCGACCAATCGCATGCGTGG 1484  
 QY 970 CCCTCCCTCCCACTCCCTCTGTGTATAAATTTACATGTGATGTCTGTGAATGCAC 1029  
 Db 1485 CCCTCCCTCCCACTCCCTCTGTGTATAAATTTACATGTGATGTCTGTGAATGCAC 1544  
 QY 1030 AAGCTAAGAGAGCTTGCAGAAAAA 1061  
 Db 1545 AAGCTAAGAGAGCTTGCAGAAAAA 1576

RESULT 12  
 ABK12143/C  
 ID ABK12143 standard; cDNA; 693 BP.  
 XX  
 AC ABK12143;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Human MIVR-1 homologous sequence #1.  
 XX  
 KW Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;  
 KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;  
 KW cardiac cell; anti-apoptotic; vascular endothelial cell;  
 KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;  
 KW heart failure; A1761441.1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200216416-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 21-AUG-2001; 2001WO-US26089.  
 XX  
 PR 22-AUG-2000; 2000US-227159P.  
 XX  
 PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
 PA (PFIZ ) PFIZER INC.  
 XX  
 PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;  
 XX  
 DR WPI; 2002-280912/32.  
 XX  
 PT Novel nucleic acid molecule encoding Mechanically Induced Vascular  
 PT Receptor-1 polypeptide, useful for treating cardiovascular diseases -  
 XX  
 PS Disclosure; Page 101; 105pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule encoding a  
 CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having  
 CC cardiac cell anti-apoptotic activity and fragments of it provided  
 CC they are not identical to Genbank sequences A1761441.1, A1594390,



QY 202 ACGGTCCTTCATCAGCCGACAGCCAGGGCGGAGAGAGATGCCCTGTCTCAGA 261  
DB 214 CGCTCTCTTCATCAGCCGACAGCCAGGCGAGGAGAGATGACTGTCTCGA 273  
QY 262 AGGATGCTGTGCGGCTCGGAGACACAGTGTGAGCAACGGAATCCAGAGCCGAGGT 321  
DB 274 AGGATGCTGTGCGGCTCGGAGACACAGTGTGAGCAACGGAATCCAGAGCCGAGGT 330  
QY 322 CTACGCGCGGCTCGGCGCCACCGAGCGCTGCGCGCTGCGCGCTGCGCGCGGAGG 381  
DB 331 CTATCGCGCGGCTCGGCGCCACCGAGCGCTGCGCGCTGCGCGCTGCGCGCGGAGG 385  
QY 382 CTTCACGCGCTTCAGCGCCACCTATCTCGTACCTGACGACGAGATCGACTCGCGCCAC 441  
DB 386 ----AGCGATTCGAACCCACCTACCTACCTGACGACGAGAAATGCGCTGCGACCCAC 441  
QY 442 CATCTCGTGTGAGCGGGAGGAGCCCCACCTTACCAAGGCGCCCTGCACCTCCAGGT 501  
DB 442 CATCTCACTGTGTGATGGGGAGGAGCCCCACCTTACCAAGGCGCCCTGCACCTCCAGGT 501  
QY 502 TCGGACCCCGAGCAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCAACAG 561  
DB 502 ACGGACCCCTGAGCAACAGCTGGAGCTGAACCGGGAATCTGTGCGCGCACCCCTTAACGG 561  
QY 562 AACCATCTTCGACAGTGAACCTGATGATGATGCTGCGGCGGCGGCGGCGGCGGCGG 621  
DB 562 GACCATCTTCGACAGTGAACCTGATGATGATGATGCTGCGGCGGCGGCGGCGGCGG 621  
QY 622 CAGTAATCTGGGCATCAGCGCCACCTGCTACGCGAGCGGCGGCGGCGGCGGCGGCGG 681  
DB 622 CAGTAATCTGGGCATCAGCGCCACCTGCTACGCGAGCGGCGGCGGCGGCGGCGGCGG 681  
QY 682 GCGGACCTACGCGAGGTGATCGGCGCATACCGGGGCTGCTCTTCAGCAGCAGAG 741  
DB 682 CCGGACCTACGCGAGGTGATCGGCGCATACCGGGGCTGCTCTTCAGCAGCAGAG 741  
QY 742 CAGTGGGCGGCTCTCTGCTGGAGGAGGAGCGGCTCCACACACACATCGCGGCGGCT 801  
DB 742 CAGTGGGCGGCTCTCTGCTGGAGGAGGAGCGGCTCCACACACACATCGCGGCGGCT 801  
QY 802 AGAGAGCGGCGCATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 861  
DB 802 GGA-----GAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 861  
QY 862 TCCGAGGGGGGCGG 876  
DB 847 GTGGGGCGGCGGCGG 861

RESULT 14  
AAA75152  
ID AAA75152 standard; cDNA; 1713 BP.

AC AAA75152;

DT 15-JAN-2001 (first entry)

XX cDNA encoding a murine TANGO 261 polypeptide.

QY TANGO 266; TANGO 216; TANGO 261; TANGO 262; TANGO 267;  
KW cellular proliferation; cellular differentiation; cellular adhesion;  
KW von Willebrand factor-associated disorder; cell trafficking; cancer;  
KW hematopoietic associated disease; atelectasis; pulmonary congestion;  
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;  
KW intestinal disorder; spleen associated disease; renal disorder;  
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;  
KW brain herniation; iatrogenic disease; inflammation; meningitis;  
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;  
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS. Mus sp.

XX Key Location/Qualifiers

XX Key

FT CDS 2..655  
FT /\*tag= a  
FT /product= "TANGO 261"  
XX  
PN WO200052022-A1.  
PD 08-SEP-2000.  
XX  
PF 01-MAR-2000; 2000WO-US05226.  
XX  
PR 01-MAR-1999; 99US-0122458.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;  
XX  
XX WPI; 2000-579269/54.  
DR P-PSDB; AAB18450.  
XX  
PT Novel human and murine secreted proteins designated TANGO 216, 261,  
PT 262, 266 and 267 useful as modulating agents of cellular processes,  
PT e.g. for treating cancer -  
XX  
PS Claim 2; Fig 6A-B; 175pp; English.  
XX  
XX The present sequence encodes a murine TANGO 261 polypeptide. The  
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and  
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular  
CC proliferation, modulate cellular differentiation and/or modulate  
CC cellular adhesion. The proteins can be used to treat any von Willebrand  
CC factor-associated disorder, regulate extracellular matrix structuring,  
CC cellular adhesion, and cell trafficking and/or migration, modulate  
CC cellular interactions, modulate cell adhesion in proliferative  
CC disorders, such as cancer, modulate the proliferation, differentiation,  
CC and/or function of cells that appear in the bone marrow, and leukocytes,  
CC treat bone marrow, blood and hematopoietic associated diseases and  
CC disorders, atelectasis, pulmonary congestion or oedema, emphysema,  
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal  
CC disorders, spleen associated diseases, modulate renal disorders, treat  
CC cardiovascular disorders such as ischemic heart disease, modulate the  
CC proliferation, differentiation, and/or function of bone and cartilage  
CC cells and to treat bone and/or cartilage associated diseases or  
CC disorder. They may also be used to treat disorders associated with the  
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic  
CC disease, inflammation, bacterial and viral meningitis, Alzheimer's  
CC disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,  
CC brain cancers, hydrocephalus and encephalitis, and treat hepatic  
CC disorders.  
XX  
SQ Sequence 1713 BP; 506 A; 438 C; 400 G; 369 T; 0 other;

Query Match 49.1%; Score 521.4; DB 21; Length 1713;  
Best Local Similarity 78.5%; Pred. No. 2.5e-103;  
Matches 704; Conservative 0; Mismatches 151; Indels 42; Gaps 5;

QY 178 GCTGAGCCACTACAGCTGTCTGCAAGCTCTTCTATCAGCGGACACGCCAGGCGGAG 237  
DB 1 GCTGAGCCACTACAGCTGTCTGAGCGGCTCTTCTATCAGCGGACACGCCAGGCGGAG 60  
QY 238 GAGAGAGATGCGCTGTCTCTCAGAGGATGCTGTGGCGCTCGGAGAGCACAGTGTCTCAGG 297  
DB 61 GAGAGACATGAGCTGTCTCTGGAAGGATGCTGTGGCGCTCAGAGAGTACGCTGTCTCAGG 120  
QY 298 CAACGGAATCCAGAGCGCGAGTCTACGCGCGCTCGGCGGACACGCCAGGCGGCGGT 357  
DB 121 --TGAATGCGGAGCCACAGGTCTATGCGCGGCTCGGCGGCTCGGCGGCTCGGCTGT 177  
QY 358 GCGGCGCTTCGCGGAGGCGGCTTCCAGCGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGG 417  
DB 178 GCGGCGCTTCGCGGAGGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGG 228  
QY 418 GCACGAGATCGACCTGCGCGGCGGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGGCTTCCAGCGG 477

```
Db 229 GCACGAATATGCCCTGCCACCCACCATCTCTACTGTCTGATGGGAGGAGCCCCACCCCTA 288
QY 478 CCAGGGCCCTTGCACCCCTCCAGCTTCGGAGACCCGAGCAGCAGCTGGAACTGAACCGGGA 537
Db 289 CCAGGGCCCTTGCACCCCTCCAGCTACGGGAGCCCTGAGCAACAGCTGGAGCTGAACCGGGA 348
QY 538 CTCGGTGGCGGACCCCAACAGAACATCTTCACAGTGACCTGATGATAGTGCCAG 597
Db 349 ATCTGTGGCGGACCCCTTAACCGGACCATCTTCGACAGTGACCTTATAGACAGCACCAT 408
QY 598 GCTGGCGGCGCCCTGCCCTCCAGCAGTAACCTCGGGCATACGGCCACCTGCTAGCGGAG 657
Db 409 GCTGGGGGCGCCCTGCTCCCGCAGCAGTAACCTCGGGCATACGGCCACCTGCTAGCAG 468
QY 658 CGCGGGCGCATGAGAGGGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGG 717
Db 469 CGGTGGGCGCATGAGAGGGCGCCGCCACCTACAGCGAGGTCTATGGCCACTACCCCTGG 528
QY 718 GTCCTCTTCAGACACACGAGCAGTGGGCGCCCTCTGCTGGAGGGACCCGGCT 777
Db 529 CTCCTCTTCAGACACACGAGCAAGTAACGGGCGCATCTCTCCCTGCTAGAGGGAGCCGGCT 588
QY 778 CCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTCGAGCAAAAGAGAGTAA 837
Db 589 CCATCTCTGCACATGCCCCACTGGA-----GAACAAGGAGAGAGAA 633
QY 838 ACAGAAAGGACACCTCTTAGGTGCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAA 897
Db 634 ACAGAAAGGTCAACCCCTTAGAGTGGGGCGGGCGCTGTAGGCAAAACCGCAAAA 693
QY 898 AAGGCAG-----AACACTCGCGCTTCTTAGAGAGAGTGTAGAGGAAGCGG 944
Db 694 AAAAAAGAAAGAAAAAACAACACTCGCGACTTCTTAAGAGAGAGAGAGAGGAAGTCA 753
QY 945 GGGGCGGAGCAACGATCTGTGGCCCTCCCTCCACCTCCCTGCTGTATATAATTTA 1004
Db 754 GGGGACACA--CAGCTGAGTGGCGGTGTGTAGTCTCTCTGTGTATATAATTTA 811
QY 1005 CATGTGATGTCTGTGCTGAATSCAAGCTTAAGAGAGCTTGCAAAAAAAGAAAAA 1061
Db 812 CATGTCTGTGTGGTCTGAATGACAGAGCTCAAAAGCTTGCAAAAAAGAAAAAAGA 868

RESULT 15
ID AAA75166 standard; cDNA; 1713 BP.
AC AAA75166;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA clone encoding a murine TANGO 261 polypeptide.
XX
KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasia;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 2..655
FT /tag= a
FT /product= "TANGO 261"
XX
PN WO200052022-AL.
XX
```

```
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 200WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX
WPI; 2000-579269/54.
DR P-PSDB; AAB18464.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
PS Disclosure; Page -; 175pp; English.
XX
CC AAA75166-68 encode murine TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells, and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
SQ Sequence 1713 BP; 505 A; 439 C; 400 G; 369 T; 0 other;

Query Match 49.0%; Score 519.8; DB 21; Length 1713;
Best Local Similarity 78.4%; Pred. No. 5.5e-103;
Matches 703; Conservative 0; Mismatches 152; Indels 42; Gaps 5;

QY 178 GCTGAGCCACTACAAAGCTGTCTGCACGGTCTTTCATCAGCGCGGCACAGCCAGGGCGGAG 237
Db 1 GCTGAGCCACTACAAAGCTGTCTGCACGGTCTTTCATCAGCGCGGCACAGCCAGGGCGGAG 60
QY 238 GAGAGAGATGCCCTGTCTCCTCAGAGGATGCTGTGGCCCTCGGAGAGACAGTGTCTAGG 297
Db 61 GAGAGACGATGCACTGTCTCCGAGGATGCTCTGGCCCTCAGAGAGTACGGTGTCTAGG 120
QY 298 CAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCTCGGCCACCGACCGCTCGCCGT 357
Db 121 ---TGGAATCGCGAGCCACAGGTCTATGCCCGCGCTCGGCCCACTGACCGACTCGCTGT 177
QY 358 GCGGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGGCCCACTTCCGTACTCTGCA 417
Db 178 GCGGCCCTTCATCAGCGG-----AGCGGATTCACCCCACTTACCCCTACTCTGCA 228
QY 418 GCACGAGATCGACCTTGGCCGCCACCATCTCTGTCAGACGGGGAGGAGCCCCACCCCTA 477
Db 229 GCAGGAATTGCCCTTGCCACCCACCATCTCTACTGTCTGATGGGAGGAGCCCCACCCCTA 288
QY 478 CCAGGGCCCTTGCACCCCTTCCAGCTTCGGGAGACCCCGAGCAGCAGCTGGAACTGAACCGGGA 537
Db 289 CCAGGGCCCTTGCACCCCTTCCAGCTACGGGAGCCCTGAGCAACAGCTGGAGCTGAACCGGGA 348
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 21:21:44 ; Search time 4130 Seconds  
(without alignments)  
10509.713 Million cell updates/sec

Title: US-09-857-826B-44

Perfect score: 1061

Sequence: 1 tcttccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rtd.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1034.4	97.5	1141	9	AF224278	AF224278 Homo sapi
2	1032.2	97.3	1818	9	AY128643	AY128643 Homo sapi
3	957.6	90.3	1061	9	BC015918	BC015918 Homo sapi
4	956	90.1	4839	9	AF305616	AF305616 Homo sapi
5	800.6	75.5	1321	6	AX392417	AX392417 Sequence
6	752.6	70.9	861	6	AX392419	AX392419 Sequence
7	748.2	70.5	61505	9	AF305426	AF305426 Homo sapi
8	748.2	70.5	130435	9	HS71877	AL035541 Human DNA
9	728.8	68.7	1583	6	AX593655	AX593655 Sequence
10	644.2	60.7	693	6	AX392430	AX392430 Sequence
11	584	55.0	1379	10	BC036995	BC036995 Mus muscu
12	523.8	49.4	878	6	AX392428	AX392428 Sequence
13	433.4	40.8	156698	10	AL837509	AL837509 Mouse DNA
14	433.4	40.8	175754	2	AC110189	AC110189 Mus muscu
15	433.4	40.8	176821	2	AL837520	AL837520 Mus muscu
16	423.4	39.9	651	10	AF220208	AF220208 Mus muscu
17	421.4	39.7	249554	2	AC139417	AC139417 Rattus no
18	421.4	39.7	258632	2	AC111878	AC111878 Rattus no
19	399.8	37.7	408	6	AX071267	AX071267 Sequence
20	399	37.6	673	6	AX525744	AX525744 Sequence
21	365.6	34.5	812	6	AX011709	AX011709 Sequence
22	350.8	33.1	2170	6	AX713513	AX713513 Sequence
23	350.8	33.1	2170	9	AK055028	AK055028 Homo sapi
24	350	33.0	8093	6	AX392432	AX392432 Sequence
25	350	33.0	8093	9	AF009426	AF009426 Homo sapi
26	349	32.9	921	6	AR233384	AR233384 Sequence
27	349	32.9	8065	6	AF233383	AF233383 Sequence
28	349	32.9	8494	9	AF009424	AF009424 Homo sapi
29	282	26.6	8039	9	AF009427	AF009427 Homo sapi
30	281	26.5	867	6	AR233385	AR233385 Sequence
31	281	26.5	1573	9	BC030199	BC030199 Homo sapi
32	281	26.5	8440	9	AF009425	AF009425 Homo sapi
33	250.2	23.6	172692	2	AF001013	AF001013 Homo sapi
34	250.2	23.6	173709	9	AF001010	AF001010 Homo sapi
35	250.2	23.6	181083	2	AP001268	AP001268 Homo sapi
36	249.2	23.5	475	6	AX392431	AX392431 Sequence
37	240.8	22.7	766	10	BC022716	BC022716 Mus muscu
38	232.6	21.9	240453	2	AC097603	AC097603 Rattus no
39	232.6	21.9	280575	2	AC117364	AC117364 Rattus no
40	231	21.8	183681	2	AC111069	AC111069 Mus muscu
41	223.8	21.1	167489	2	BX005347	BX005347 Danio rer
42	206.4	19.5	176458	2	AC134911	AC134911 Mus muscu
43	182.4	17.2	155348	5	AL928820	AL928820 Zebrafish
44	158.6	14.9	150224	9	HSJ1059L7	AL121913 Human DNA
45	103.6	9.8	2570	9	AK056098	AK056098 Homo sapi

#### ALIGNMENTS

RESULT 1  
AF224278  
LOCUS AF224278 Homo sapiens PMEPA1 protein (PMEPA1) mRNA, complete cds.  
DEFINITION AF224278  
ACCESSION AF224278  
VERSION AF224278.1 GI:9255808  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1141)  
AUTHORS Xu,L.L., Sharmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,  
Moul,J.W. and Srivastava,S.  
TITLE A novel androgen-regulated gene, PMEPA1, located on chromosome

20q13 exhibits high level expression in prostate  
Genomics 66 (3), 257-263 (2000)

OLINE 20334621

JBMED 10873380

REFERENCE  
100,000  
2 (base 1 to 1141)

2 (Bases 1 to 1141)  
Xu, L.L., Shanmugam, N., Segawa, T., Sesterhenn, I.A., McLeod, D.G.,  
Moul, J.W. and Srivastava, S.

# FILE

JOURNAL,

Submitted (12-04V-2007) CFDR, USGMS, 1330 East Vellefson Street,  
Rockville, MD 20852, USA  
Registration/Qualification

**UKES**

**2008**

```
/mol_type="mRNA"
```

/db\_xref:

/chromosome="20"

```
/var/www/html/2013-11-13-33
```

1100-1110-1120-1130-1140-1150-1160-1170-1180-1190-1200-1210-1220-1230-1240-1250-1260-1270-1280-1290-1300-1310-1320-1330-1340-1350-1360-1370-1380-1390-1400-1410-1420-1430-1440-1450-1460-1470-1480-1490-1500-1510-1520-1530-1540-1550-1560-1570-1580-1590-1600-1610-1620-1630-1640-1650-1660-1670-1680-1690-1700-1710-1720-1730-1740-1750-1760-1770-1780-1790-1800-1810-1820-1830-1840-1850-1860-1870-1880-1890-1900-1910-1920-1930-1940-1950-1960-1970-1980-1990-2000-2010-2020-2030-2040-2050-2060-2070-2080-2090-2100-2110-2120-2130-2140-2150-2160-2170-2180-2190-2200-2210-2220-2230-2240-2250-2260-2270-2280-2290-2300-2310-2320-2330-2340-2350-2360-2370-2380-2390-2400-2410-2420-2430-2440-2450-2460-2470-2480-2490-2500-2510-2520-2530-2540-2550-2560-2570-2580-2590-2600-2610-2620-2630-2640-2650-2660-2670-2680-2690-2700-2710-2720-2730-2740-2750-2760-2770-2780-2790-2800-2810-2820-2830-2840-2850-2860-2870-2880-2890-2900-2910-2920-2930-2940-2950-2960-2970-2980-2990-3000-3010-3020-3030-3040-3050-3060-3070-3080-3090-3100-3110-3120-3130-3140-3150-3160-3170-3180-3190-3200-3210-3220-3230-3240-3250-3260-3270-3280-3290-3300-3310-3320-3330-3340-3350-3360-3370-3380-3390-3400-3410-3420-3430-3440-3450-3460-3470-3480-3490-3500-3510-3520-3530-3540-3550-3560-3570-3580-3590-3600-3610-3620-3630-3640-3650-3660-3670-3680-3690-3700-3710-3720-3730-3740-3750-3760-3770-3780-3790-3800-3810-3820-3830-3840-3850-3860-3870-3880-3890-3900-3910-3920-3930-3940-3950-3960-3970-3980-3990-4000-4010-4020-4030-4040-4050-4060-4070-4080-4090-4100-4110-4120-4130-4140-4150-4160-4170-4180-4190-4200-4210-4220-4230-4240-4250-4260-4270-4280-4290-4300-4310-4320-4330-4340-4350-4360-4370-4380-4390-4400-4410-4420-4430-4440-4450-4460-4470-4480-4490-4500-4510-4520-4530-4540-4550-4560-4570-4580-4590-4600-4610-4620-4630-4640-4650-4660-4670-4680-4690-4700-4710-4720-4730-4740-4750-4760-4770-4780-4790-4800-4810-4820-4830-4840-4850-4860-4870-4880-4890-4900-4910-4920-4930-4940-4950-4960-4970-4980-4990-5000-5010-5020-5030-5040-5050-5060-5070-5080-5090-5100-5110-5120-5130-5140-5150-5160-5170-5180-5190-5200-5210-5220-5230-5240-5250-5260-5270-5280-5290-5300-5310-5320-5330-5340-5350-5360-5370-5380-5390-5400-5410-5420-5430-5440-5450-5460-5470-5480-5490-5500-5510-5520-5530-5540-5550-5560-5570-5580-5590-5600-5610-5620-5630-5640-5650-5660-5670-5680-5690-5700-5710-5720-5730-5740-5750-5760-5770-5780-5790-5800-5810-5820-5830-5840-5850-5860-5870-5880-5890-5900-5910-5920-5930-5940-5950-5960-5970-5980-5990-6000-6010-6020-6030-6040-6050-6060-6070-6080-6090-6100-6110-6120-6130-6140-6150-6160-6170-6180-6190-6200-6210-6220-6230-6240-6250-6260-6270-6280-6290-6300-6310-6320-6330-6340-6350-6360-6370-6380-6390-6400-6410-6420-6430-6440-6450-6460-6470-6480-6490-6500-6510-6520-6530-6540-6550-6560-6570-6580-6590-6600-6610-6620-6630-6640-6650-6660-6670-6680-6690-6700-6710-6720-6730-6740-6750-6760-6770-6780-6790-6800-6810-6820-6830-6840-6850-6860-6870-6880-6890-6900-6910-6920-6930-6940-6950-6960-6970-6980-6990-7000-7010-7020-7030-7040-7050-7060-7070-7080-7090-7100-7110-7120-7130-7140-7150-7160-7170-7180-7190-7200-7210-7220-7230-7240-7250-7260-7270-7280-7290-7300-7310-7320-7330-7340-7350-7360-7370-7380-7390-7400-7410-7420-7430-7440-7450-7460-7470-7480-7490-7500-7510-7520-7530-7540-7550-7560-7570-7580-7590-7600-7610-7620-7630-7640-7650-7660-7670-7680-7690-7700-7710-7720-7730-7740-7750-7760-7770-7780-7790-7800-7810-7820-7830-7840-7850-7860-7870-7880-7890-7900-7910-7920-7930-7940-7950-7960-7970-7980-7990-8000-8010-8020-8030-8040-8050-8060-8070-8080-8090-8100-8110-8120-8130-8140-8150-8160-8170-8180-8190-8200-8210-8220-8230-8240-8250-8260-8270-8280-8290-8300-8310-8320-8330-8340-8350-8360-8370-8380-8390-8400-8410-8420-8430-8440-8450-8460-8470-8480-8490-8500-8510-8520-8530-8540-8550-8560-8570-8580-8590-8600-8610-8620-8630-8640-8650-8660-8670-8680-8690-8700-8710-8720-8730-8740-8750-8760-8770-8780-8790-8800-8810-8820-8830-8840-8850-8860-8870-8880-8890-8900-8910-8920-8930-8940-8950-8960-8970-8980-8990-9000-9010-9020-9030-9040-9050-9060-9070-9080-9090-9100-9110-9120-9130-9140-9150-9160-9170-9180-9190-9200-9210-9220-9230-9240-9250-9260-9270-9280

```
1 1743 /cell_line="LNCaP"
```

gene  
1. 1141

```
/gene="PMEPA1"
```

CDS 96.854

```
/gene="PMEPA1"
```

/note="t

induced in res

induced in response to the synthetic androgen, R1881; expression is abundant in, and restricted to, prostate glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP 004329"

/codon start=1

```

"product" : "LVBEBB1",
"category" : "T-shirts"

```

```
product=PMERAT protein/
truncat=13 truncat=13
```

```
/protein_id="AAF86322
```

```

/db_xref="GI:9255809"
/translation="MARLEFVQIIIIIVVMMVVITLLSHYKLSARSFIRHSQG
RRRDLASSEGLMPSTVSGNGIPQVYAPPRDTRLVAPFPAQRERFHRFQTY
PYLQHEIDLPTISLDGEPPPPVQGECTIQLRDPEQQLNRESVAPNRITFUSD
LMDSARGLPPSPSGNSGIGATCVSGRMEGQPPPTVSEVIGHPGSSFOHQSSGPP
SLIRGTLHHTHTIATPISSAAIWSKSEKOKKGGHDI"

```

COUNT	271	3	350	336	3	184
	CCECOTINEHIIHRE DESBATAWSNEK					

187

dry Match	97.5†	Score 1034.4;	DB 9;	Length 1141;
at Local Similarity	99.2†	Pred. No. 9.2e-185;		
ches 1050: Conservative	0:	Mismatches	6:	Indels
			2:	Gaps
				1:

4 TCCTTGGGTTCCGGGTGAAAGCGCTTGGGGTTCAGTGGGCCATGATCCCCGAGTGTCTGG 63  
|||||  
1 TCCTTGGGTTCCGGGTGAAAGCGCTGGGGTTC - GTGGCCATGATCCCGAGTGTCTGG 58  
|||||  
64 AGAACTGAAGCGGACGGTCTCTCGGAAACAGGCGAATGGCGGAGCTGGAGTTTGTTC 123  
|||||  
59 AGAACTGAAGCGGACAGTCTCTCGGAAACAGGCGAATGGCGGAGCTGGAGTTTGTTC 118  
|||||  
124 GATCATCATCATCTGTGTGTGATGATGTGTGGTGGTGATCAGTGCCTGCTGAG 183  
|||||  
119 GATCATCATCATCTGTGTGTGATGATGTGTGGTGGTGATCAGTGCCTGCTGAG 178  
|||||  
184 CCACCTCAAGGTGTCGACGGTCTTTCATCAGCCGCGACAGCCAGGGCGGAGGAGAGA 243  
|||||  
179 CCACCTCAAGGTGTCTGACGGTCTTTCATCAGCCGCGACAGCCAGGGCGGAGGAGAGA 238  
|||||  
244 AGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGCAGGCAACGG 303  
|||||  
239 AGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGCAGGCAACGG 298  
|||||  
304 AATCCGAGAGCGCAGGTCTACGCCCGCCTCGGCCACCGACCGCTGGCCGTGCGGCC 363  
|||||  
299 AATCCGAGCGCAGGTCTACGCCCGCCTCGGCCACCGACCGCTGGCCGTGCGGCC 358  
|||||  
364 CTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGAGCACGA 423  
|||||  
359 CTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGAGCACGA 418  
|||||  
424 GATCGACTGCGGCCCAACATCTCTGCTGTGACAGCGGGGAGAGCCCCCACTTACCGAGG 483

FEATURES	Location/Qualifiers	
	1..1818	
source	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/chromosome="20"	
	/map="20q13.31-13.33"	
	145..858	
CDS	/codon_start=1	
	/product="PMEPA1 variant A protein"	
	/protein_id="AA089277.1"	
	/db_xref="GI:22121999"	
	/translation="MMWVVTICLLSHYKLSARFISRHSGRRRDALSSRGCLWP	
	SESTVSGNGIPEPQYAPPRDLAVPPFAQRERFHFQPTYPLQHEIDLPTISL	
	SDGEPYPYQPFQLQLRDEPQQLELNRESVRAPNRTIFDSLDMSARLGGCPSPS	
	NSGTSATCYGSGRMGPPPTYSVIGHYPGSSFFHQHQSSGPPSLLEGLRLHHTIAP	
	LESAIWSKXDKQKHPL"	
	459 a 467 c 494 g 398 t	
BASE COUNT	459 a 467 c 494 g 398 t	
ORIGIN		
Query Match	97.3%; Score 1032.2; DB 9; Length 1818;	
Best Local Similarity	99.1%; Pred. No. 2.3e-184;	
Matches 1051; Conservative	0; Mismatches 3; Indels 7; Gaps 1;	
QY	1	TCCTCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
Db	7	TCCTCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 66
QY	61	TGGAGAACTGAAGCGGAGCGGTCTCCTCGGAAACCCAGGCAATGGCGGAGCTGGAGTTGT 120
Db	67	TGGAGAACTGAAGCGGAGCAGTCTCTCTCGGAAACCA-----GCGGAGCTGGAGTTGT 119
QY	121	TCAGATCATCATCATCGTGGTGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Db	120	TCAGATCATCATCATCGTGGTGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGT 179
QY	181	GAGCACTACAAGCTGTGTGACAGCTCTTATCAGCGCGCACACCCAGGCGGAGGAG 240
Db	180	GAGCACTACAAGCTGTGTGACAGCTCTTATCAGCGCGCACACCCAGGCGGAGGAG 239
QY	241	AGAAGATGCCCTGTCTCAGAGGAGTGCCTGTGGCCCTCGGAGACAGAGTGTACAGCAA 300
Db	240	AGAAGATGCCCTGTCTCAGAGGAGTGCCTGTGGCCCTCGGAGACAGAGTGTACAGCAA 299
QY	301	CGGAATCCAGACCGCGAGGCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGCC 360
Db	300	CGGAATCCAGACCGCGAGGCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGCC 359
QY	361	GCCCTTCGCCCCAGCGGAGCGCTTCCAGCGTTCCAGGCCACCTATCCGTACCTGCAGCA 420
Db	360	GCCCTTCGCCCCAGCGGAGCGCTTCCAGCGTTCCAGGCCACCTATCCGTACCTGCAGCA 419
QY	421	CGAGATCGACCTGCCGCCACCATCTCGTGTACAGCGGGAGAGCCCCCACCCTACCA 480
Db	420	CGAGATCGACCTGCCGCCACCATCTCGTGTACAGCGGGAGAGCCCCCACCCTACCA 479
QY	481	GGGCCCCCTGCACCTTCAGTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTC 540
Db	480	GGGCCCCCTGCACCTTCAGTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTC 539
QY	541	GGTGCGCGACCCCCAAACAGAACCATCTTCGACAGTGCACCTGTAGTGGATGCCAGGCT 600
Db	540	GGTGCGCGACCCCCAAACAGAACCATCTTCGACAGTGCACCTGTAGTGGATGCCAGGCT 599
QY	601	GGGCGGCCCCCTGCCCGCCAGCAGTAATCTCGGCGATCAGCGCACGTGCTACGGCAGCG 660
Db	600	GGGCGGCCCCCTGCCCGCCAGCAGTAATCTCGGCGATCAGCGCACGTGCTACGGCAGCG 659
QY	661	CGGCGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTTCATCGGCCCTACCCGGGGTC 720
Db	660	CGGCGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTTCATCGGCCCTACCCGGGGTC 719
QY	721	CTCCTTTCCAGCACCGAGCAGAGTGGGCGCGCCCTCTTGTCTGGAGGGGAGCCCGGCTCCA 780
Db	720	CTCCTTTCCAGCACCGAGCAGAGTGGGCGCGCCCTCTTGTCTGGAGGGGAGCCCGGCTCCA 779
QY	781	CCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 840
Db	780	CCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 839
QY	841	GAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGTGGTGAAG 900
Db	840	GAAGAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGGTGGTGAAG 899
QY	901	GCAGAACCTCCGCGCTCTTAGAGAGGAGTGTAGAGGAAGCGCGGGGGCGCAACGC 960
Db	900	GCAGAACCTCCGCGCTCTTAGAGAGGAGTGTAGAGGAAGCGCGGGGGCGCAACGC 959
QY	961	ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTGTGTC 1020
Db	960	ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTGTGTC 1019
QY	1021	TGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 1061
Db	1020	TGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 1060
RESULT 3		
LOCUS	BC015918 1061 bp mRNA linear PRI 17-OCT-2001	
DEFINITION	Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.	
ACCESSION	BC015918	
VERSION	BC015918.1 GI:16198474	
KEYWORDS	MGC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 1061)	
JOURNAL	Strausberg,R.	
	Direct Submission	
	Submitted (15-OCT-2001) National Institutes of Health, Mammalian	
	Gene Collection (MGC), Cancer Genomics Office, National Cancer	
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
	USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk	
	Email: gcgaps-r@mail.nih.gov	
	Tissue Procurement: DCTD/BTP	
	cDNA Library Preparation: Rubin Laboratory	
	DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)	
	BC Cancer Agency, Vancouver, BC, Canada	
	info@bcgsc.bc.ca	
	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,	
	Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,	
	Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo	
	Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven	
	Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline	
	Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,	
	Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,	
	George Yang, Scott Zuyderduyn, Marco Marra.	
Clone distribution: MGC clone distribution information can be found		
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Series: IRAL Plate: 29 Row: e Column: 5		
This clone was selected for full length sequencing because it		
passed the following selection criteria: Similarity but not		
identity to protein.		
FEATURES	Location/Qualifiers	
	1..1061	
source	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="MGC:20374 IMAGE:4559576"	
	/tissue type="Kidney, renal cell adenocarcinoma"	



Qy	142	GGTGATGATGGTGTGGTGGTGTGATCAGCTGCTGCTGATGAGCCACTACAAGCTGTCTGC	201
Db	467	GGTGATGATGGTGTGGTGGTGTGATCAGCTGCTGCTGATGAGCCACTACAAGCTGTCTGC	526
Qy	202	ACGGTCTCTTATCAGCCGGCAGCAGCGGGGAGGAGAGAGATGCCCTGTCTCAGA	261
Db	527	ACGGTCTCTTATCAGCCGGCAGCAGCGGGGAGGAGAGAGATGCCCTGTCTCAGA	586
Qy	262	AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGCAACGGAATCCAGAGCGCAGGT	321
Db	587	AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGCAACGGAATCCAGAGCGCAGGT	646
Qy	322	CTAGCCCGGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTTGGCGGCGAGGAGCG	381
Db	647	CTAGCCCGGCTCGGCCACCGACCGCTGCGCGCTGCGCGCTTGGCGGCGAGGAGCG	706
Qy	382	CTTCCACCGCTTCAGAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCGGCCAC	441
Db	707	CTTCCACCGCTTCAGAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCGGCCAC	766
Qy	442	CATCTCGTGTGACAGCGGGAGGACCCCACTACCTACAGGCGCCCTGCACCTCCAGCT	501
Db	767	CATCTCGTGTGACAGCGGGAGGACCCCACTACCTACAGGCGCCCTGCACCTCCAGCT	826
Qy	502	TCGGGACCCGAGCAGCAGCTGGAACCTGAAACCGGAGTGGTGGCGGACACCCCAACAG	561
Db	827	TCGGGACCCGAGCAGCAGCTGGAACCTGAAACCGGAGTGGTGGCGGACACCCCAACAG	886
Qy	562	AACCATCTTCAGACGTGACCTGATGATAGTGCAGGCTGGGCGCCCTGCCGCCCCAG	621
Db	887	AACCATCTTCAGACGTGACCTGATGATAGTGCAGGCTGGGCGCCCTGCCGCCCCAG	946
Qy	622	CAGTAACCTCGGCACTCAGCGGCACTGCTGCTACGCGCGCGGCGCATGAGGGCGGCC	681
Db	947	CAGTAACCTCGGCACTCAGCGGCACTGCTGCTACGCGCGCGGCGCATGAGGGCGGCC	1006
Qy	682	GCCCACTTACAGCAGGATCATCGGCCACTACCGGGGCTCTCTTCCAGCACACAGCAG	741
Db	1007	GCCCACTTACAGCAGGATCATCGGCCACTACCGGGGCTCTCTTCCAGCACACAGCAG	1066
Qy	742	CAGTGGGCGCCCTCTCTTGGAGGGACCGGCTCCACACACACACATCGCGCCCT	801
Db	1067	CAGTGGGCGCCCTCTCTTGGAGGGACCGGCTCCACACACACACATCGCGCCCT	1126
Qy	802	AGAGAGCGCAGCCTCTCGAGCAAGAGAGATAAACAGAAAGACACCTCTCTAGGG	861
Db	1127	AGAGAGCGCAGCCTCTCGAGCAAGAGAGATAAACAGAAAGACACCTCTCTAGGG	1186
Qy	862	TCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAAGGCGAGAACACTCCCGCTTCTT	921
Db	1187	TCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAAGGCGAGAACACTCCCGCTTCTT	1246
Qy	922	AGAGAGAGTGAGAGGAGGGGGGCGGAGCAACCATCGTGGCGCCCTCCCTCC	981
Db	1247	AGAGAGAGTGAGAGGAGGGGGGCGGAGCAACCATCGTGGCGCCCTCCCTCC	1306
Qy	982	ACCTCCCTGTGTATAAATATTACATGTGATGCTGTGGTCTGAATGCACAAAGTAAGAG	1041
Db	1307	ACCTCCCTGTGTATAAATATTACATGTGATGCTGTGGTCTGAATGCACAAAGTAAGAG	1366
Qy	1042	CTTGCAAAAAAAGAAAAA 1061	
Db	1367	CTTGCAAAAAAAGAAAAA 1386	

RESULT 5  
AX392417  
LOCUS AX392417 1321 bp DNA linear PAT 23-MAR-2002  
DEFINITION Sequence 1 from Patent WO0216416.  
ACCESSION AX392417  
VERSION AX392417.1 GI:19700732  
KEYWORDS Homo sapiens (human)  
SOURCE

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Lee, R.T., Landschulz, K.T., Kennedy, S.P., Thompson, J.F. and Turi, T.G.
TITLE	Diagnosis and treatment of cardiovascular conditions
JOURNAL	Patent: WO 0216416-A 1 28-FEB-2002;
FEATURES	THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
source	Location/Qualifiers
	1. 1321
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
CDS	/db_xref="taxon:9606"
	413..1276
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAD29005.1"
	/db_xref="GI:19700733"
	/translation="MRLMGVSTAAAGQPNVSTCNCKRSLFQSMETELBFLVQI
	IIIVVMVMVVITCLSHYKLSARFISRHQRRERDALSGELWPSSESTVSGN
	GIPEPVYAPRPTDLAVPPFAORFPHRFOPTYPYLOHEIDLPPTISLDGEPFP
	YQGPCTIQLRDPQOQLNRESVRAPNRTIFDSLDMSARLGPCPSSNSGTSATC
	YGSGRMEGPPPTYSYSEVIGHYPSGSPFHQSSGPPSLLEGTRLHHTHIAPLSRAIWS
	KEKDKQKHPL"
BASE COUNT	223 a 493 c 440 g 165 t
ORIGIN	
	Query Match 75.5%; Score 800.6; DB 6; Length 1321;
	Best Local Similarity 98.3%; Pred. No. 9.3e-141;
	Matches 809; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy	82 TCTCTCTCGAAACAGGCAATGCGGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 141
Db	499 TTTGTTCCAGAGCATGAGATCAGGAGCTGGAGTTGTTTCAGATCATCATCGTGGT 558
Qy	142 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 201
Db	559 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 618
Qy	202 ACGTCTCTTATCAGCGGCGACAGCGGCGGAGGAGAGAGAGAGAGAGAGAGAGAG 261
Db	619 ACGTCTCTTATCAGCGGCGACAGCGGCGGAGGAGAGAGAGAGAGAGAGAGAGAG 678
Qy	262 AGGATGCTGTGGCCCTCGGAGAGCAGTGTGAGGCAACGGAATCCAGAGCGCAGGT 321
Db	679 AGGATGCTGTGGCCCTCGGAGAGCAGTGTGAGGCAACGGAATCCAGAGCGCAGGT 738
Qy	322 CTACGCGCGGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 381
Db	739 CTACGCGCGGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 798
Qy	382 CTTCACCGCTTCAGGCCACCTATCGTACCTGACGACGAGATCGACCTGCGGCCAC 441
Db	799 CTTCACCGCTTCAGGCCACCTATCGTACCTGACGACGAGATCGACCTGCGGCCAC 858
Qy	442 CATCTCGTGTGACAGCGGAGGAGCCGCCCTTACAGGCGCCCTGCACCTCCAGCT 501
Db	859 CATCTCGTGTGACAGCGGAGGAGCCGCCCTTACAGGCGCCCTGCACCTCCAGCT 918
Qy	502 TCGGGAACCCGAGCAGCAGTGGAACTGAACCGGAGTGGTGGCGGACACCCCAACAG 561
Db	919 TCGGGAACCCGAGCAGCAGTGGAACTGAACCGGAGTGGTGGCGGACACCCCAACAG 978
Qy	562 AACCATCTTCAGACGTGACCTGATGATGATGATGATGATGATGATGATGATGATG 621
Db	979 AACCATCTTCAGACGTGACCTGATGATGATGATGATGATGATGATGATGATGATG 1038
Qy	622 CAGTAACCTCGGGCATCAGCGCCAGTGTCTAGCGGAGGCGGCGGCGCATGAGAGGGCGGCC 681
Db	1039 CAGTAACCTCGGGCATCAGCGCCAGTGTCTAGCGGAGGCGGCGGCGCATGAGAGGGCGGCC 1098
Qy	682 GCCCACCTTACAGCAGGAGGTTCATCGGCCACTACCGGGGGTCTCTCTTCCAGCACCAGCAG 741





```

repeat_region 1177. 1820
/Note="4 copies 161 mer 64% conserved"
1204. 1371
/Note="3 copies 56 mer 75% conserved"
1358. 2757
/Note="28 copies 50 mer 54% conserved"
1675. 2718
/Note="18 copies 58 mer 55% conserved"
1843. 2346
/Note="9 copies 56 mer 64% conserved"
1852. 2079
/Note="4 copies 57 mer 86% conserved"
2028. 2632
/Note="11 copies 55 mer 60% conserved"
2087. 2257
/Note="3 copies 57 mer 83% conserved"
2228. 2497
/Note="5 copies 54 mer 75% conserved"
2583. 2716
/Note="2 copies 67 mer 82% conserved"
3378. 3426
/Note="L1MB4 repeat: matches 6088. 6136 of consensus"
3799. 3896
/Note="Charliel repeat: matches 681. 781 of consensus"
5331. 5793
/Note="MLT1B repeat: matches 14. 466 of consensus"
5797. 5988
/Note="6 copies 32 mer 86% conserved"
5855. 5978
/Note="31 copies 4 mer gcac 61% conserved"
6474. 6591
/Note="MIR repeat: matches 91. 218 of consensus"
6592. 6723
/Note="FLAM C repeat: matches 1. 132 of consensus"
6952. 7021
/Note="L2 repeat: matches 2637. 2705 of consensus"
7358. 7671
/Note="AluX repeat: matches 1. 312 of consensus"
8521. 8554
/Note="Alu repeat: matches 1. 34 of consensus"
8863. 9217
/Note="MLT1D repeat: matches 105. 505 of consensus"
9228. 9746
/Note="L2 repeat: matches 1459. 2028 of consensus"
9880. 10089
/Note="L2 repeat: matches 2292. 2511 of consensus"
10094. 10206
/Note="AluJb repeat: matches 188. 300 of consensus"
10263. 10573
/Note="AluJb repeat: matches 1. 308 of consensus"
10658. 10753
/Note="L2 repeat: matches 2647. 2749 of consensus"
11816. 12380
/Note="LTR19B repeat: matches 1. 580 of consensus"
12519. 12813
/Note="L1MD2 repeat: matches 6032. 6331 of consensus"
13559. 13726
/Note="3 copies 56 mer 73% conserved"
13585. 13716
/Note="22 copies 6 mer ctctct 67% conserved"
13586. 13717
/Note="66 copies 2 mer tc 68% conserved"
13588. 13715
/Note="4 copies 32 mer 78% conserved"
13589. 13724
/Note="34 copies 4 mer cttt 77% conserved"
complement(13622. 14142)
/Note="match: GSS: Em:AQ592603"
14313. 14437
/Note="LTR16C repeat: matches 257. 386 of consensus"
15015. 15110
/Note="3 copies 32 mer 79% conserved"
15021. 15110

repeat_region 15026. 15109
/Note="7 copies 12 mer 82% conserved"
15028. 15111
/Note="14 copies 6 mer cacaca 82% conserved"
15029. 15108
/Note="20 copies 4 mer acac 83% conserved"
15273. 15399
/Note="L2 repeat: matches 2159. 2285 of consensus"
15420. 15721
/Note="AluX repeat: matches 1. 303 of consensus"
16333. 16396
/Note="2 copies 32 mer 98% conserved"
16575. 16690
/Note="L2 repeat: matches 2596. 2711 of consensus"
18049. 18169
/Note="MIR repeat: matches 86. 211 of consensus"
18312. 18438
/Note="MIR repeat: matches 138. 250 of consensus"
18585. 18776
/Note="MER20 repeat: matches 7. 217 of consensus"
19211. 19290
/Note="MIR repeat: matches 72. 154 of consensus"
20413. 20463
/Note="MIR repeat: matches 90. 143 of consensus"
complement(21005. 21672)
/Note="match: GSS: Em:AQ748384"
/Note="match: GSS: Em:AQ76209"
complement(21278. 21663)
/Note="match: GSS: Em:AQ136459"
complement(21392. 21672)
/Note="match: GSS: Em:AQ892114"
21667. 21775
/Note="MIR repeat: matches 73. 192 of consensus"
23577. 23646
/Note="2 copies 35 mer 100% conserved"
24257. 24591
/Note="L2 repeat: matches 2336. 2710 of consensus"
25580. 25611
/Note="16 copies 2 mer tc 87% conserved"
26333. 26643
/Note="AluJ repeat: matches 1. 310 of consensus"
27603. 27684
/Note="L2 repeat: matches 2661. 2739 of consensus"
27823. 28041
/Note="MIR repeat: matches 38. 242 of consensus"
28129. 28259
/Note="FLAM C repeat: matches 1. 127 of consensus"
31203. 31356
/Note="L1MB5 repeat: matches 6015. 6176 of consensus"
32224. 32522
/Note="AluX repeat: matches 1. 298 of consensus"

Query Match 70.5%; Score 748.2; DB 9; Length 130435;
Best Local Similarity 98.9%; Pred No. 5e-131;
Matches 753; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 301 CGGAATCCAGAGCGCAGGTCTACGCCCGCCCTCGGCCACCGACCGCTGGCGTGC 360
DB 128454 CTGCTTCTCCAGCGCAGGTCTACGCCCGCCCTCGGCCACCGACCGCTGGCGTGC 128395

QY 361 GCCCTTCCAGCGGAGCGCTTCACCGCTTCAGCCCGACCTATCGTACCTCAGCA 420
DB 128394 GCCCTTCCAGCGGAGCGCTTCACCGCTTCAGCCCGACCTATCGTACCTCAGCA 128335

QY 421 CGAGATGACCTCGCCGCCACCATCTCGTGTGACAGCGGAGGAGCCCGACCTACCA 480
DB 128334 CGAGATGACCTCGCCGCCACCATCTCGTGTGACAGCGGAGGAGCCCGACCTACCA 128275

QY 481 GGGCCCTGCACTTCAGCTTCGGGACCCCGAGCAGCAGTGTGAACCTGAAACCGGAGTC 540
DB 128274 GGGCCCTGCACTTCAGCTTCGGGACCCCGAGCAGCAGTGTGAACCTGAAACCGGAGTC 128215

```



QY	541	GGTGGCGGCA	CCCCAAA	CAGAA	CCATCTT	CGACAGT	GCACCTG	ATGAT	TAGTCC	CAGGCT	600	
Db	128214	GGTGGCGGCA	CCCCAAA	CAGAA	CCATCTT	CGACAGT	GCACCTG	ATGAT	TAGTCC	CAGGCT	128155	
QY	601	GGGCGGCGCT	CGCCCGC	CAGAGT	AACTC	GGGGCAT	CAGGCG	CACTG	TAGTCC	CAGGCG	660	
Db	128154	GGGCGGCGCT	CGCCCGC	CAGAGT	AACTC	GGGGCAT	CAGGCG	CACTG	TAGTCC	CAGGCG	128095	
QY	661	CGGGCGCAT	TGGAGG	CGCGCC	CCACCT	TACAGC	GAGGT	CAT	CGGCCA	CTAC	CGGGGTC	720
Db	128094	CGGGCGCAT	TGGAGG	CGCGCC	CCACCT	TACAGC	GAGGT	CAT	CGGCCA	CTAC	CGGGGTC	128035
QY	721	CTCCTTCAG	ACACG	ACAGAG	AGTGG	CGCGCC	CTCTT	CGTGG	AGGGA	CCCGG	CTCCA	780
Db	128034	CTCCTTCAG	ACACG	ACAGAG	AGTGG	CGCGCC	CTCTT	CGTGG	AGGGA	CCCGG	CTCCA	127975
QY	781	CCACACAC	ATCTCG	CGCCCT	TAGAG	CGCGC	ACATCT	CGGACA	AGAG	GAAT	TAACA	840
Db	127974	CCACACAC	ATCTCG	CGCCCT	TAGAG	CGCGC	ACATCT	CGGACA	AGAG	GAAT	TAACA	127915
QY	841	GAAGGAC	ACCTCT	CTAGG	GTCCC	CAGGG	CGCGG	CTGGG	GTGCG	TAGGT	GAAGA	900
Db	127914	GAAGGAC	ACCTCT	CTAGG	GTCCC	CAGGG	CGCGG	CTGGG	GTGCG	TAGGT	GAAGA	127855
QY	901	GCAGAAC	ACTCCG	CGCTT	CTTAG	AGAG	AGTGA	GAG	GAAG	CGCGG	CGCAG	960
Db	127854	GCAGAAC	ACTCCG	CGCTT	CTTAG	AGAG	AGTGA	GAG	GAAG	CGCGG	CGCAG	127795
QY	961	ATCGTGTG	CGCTCC	CCCTCC	ACCTCC	CTGCT	GTAT	AAAT	TATTA	TACAT	GTGTC	1020
Db	127794	ATCGTGTG	CGCTCC	CCCTCC	ACCTCC	CTGCT	GTAT	AAAT	TATTA	TACAT	GTGTC	127735
QY	1021	TGAATGCA	CAAGCT	TAAGAG	CGTTC	GCAAAA	AAAAAAAA	AAAA				1061
Db	127734	TGAATGCA	CAAGCT	TAAGAG	CGTTC	GCAAAA	AAAAAAAA	AAAA				127694
RESULT 9		AX593655				1583 bp	DNA	linear	PAT 13-FEB-2003			
LOCUS		AX593655				Sequence 32 from Patent WO0242776.						
DEFINITION		AX593655										
ACCESSION		AX593655.1				GI:28375034						
VERSION		AX593655.1										
KEYWORDS												
SOURCE												
ORGANISM						Homo sapiens (human)						
REFERENCE												
AUTHORS						Sun, Y., Recipon, H., Chen, S.Y. and Liu, C.						
TITLE						Compositions and methods relating to prostate specific genes and proteins						
JOURNAL						Patent: WO 0242776-A 32 30-MAY-2002;						
FEATURES						Location/Qualifiers						
source						1..1583						
BASE COUNT						361 a 469 c 457 g 296 t						
ORIGIN												
Query Match						68.7%; Score 728.8; DB 6; Length 1583;						
Best Local Similarity						98.9%; Pred. No. 3e-127; Indels 1; Gaps 1;						
Matches						744; Conservative 0; Mismatches 7; Indels 1; Gaps 1;						
QY	311	GAGCGCAG	GTCTAC	CGCCCG	CGCTCG	CGCCAC	CGCAG	CGCTG	CGCG	CGCTTC	CGCC	370
Db	825	GCGCGCAG	GTCTAC	CGCCCG	CGCTCG	CGCCAC	CGCAG	CGCTG	CGCG	CGCTTC	CGCC	884
QY	371	CAGCGGAG	CGCTTC	CACGCT	TTCAG	CGCCAC	CTAT	CGTAC	CTCG	CACGAG	ATCGAC	430

[illegible]

RESULT 11	
BC036995	
LOCUS	1379 bp mRNA linear ROD 26-SEP-2002
DEFINITION	Mus musculus, similar to transmembrane, prostate androgen induced RNA, clone IMAGE:5038092, mRNA.
ACCESSION	BC036995
VERSION	BC036995.1 GI:23331176
KEYWORDS	.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 1379)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (23-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NTH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 156698)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Direct Submission  
Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Nov 15, 2002 this sequence version replaced gi:24414713.  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a phred quality of at least 30.

-----Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) RP23-44L6 is  
from the RPCI-23 Mouse PAC library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.

FEATURES  
source

1..156698  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosomes="2"  
/clone="RP23-44L6"  
/clone\_lib="RPCI-23"

BASE COUNT 39443 a 38596 c 39859 g 38800 t

ORIGIN

Query Match 40.8%; Score 433.4; DB 10; Length 156698;  
Best Local Similarity 77.3%; Pred. No. 9e-72;  
Matches 593; Conservative 0; Mismatches 136; Indels 38; Gaps 4;  
QY 307 CCCAGAGCGCAGGTCTACGCCCCGCTCGGCCACCGACCGCTGGCGTGGCCGCTT 366  
Db 135333 CCTCCAGCCACAGGCTATGCTCCGCGCTCGGCCACTGACCGACTCGCTGTGCCGCCCTT 135274  
QY 367 GCGCCAGCGGAGCGCTTCCAGCGCTTCCAGCGCCACCTATCGTACCTGCACAGAGAT 426  
Db 135273 CATCCAGCGGAG-----CCGATTCACACCCACTTACCTGACGACGAAT 135223  
QY 427 CGACCTCGCCGCCACCATCTCGTGTGCAGACGGGAGGAGCCCCACCCCTACAGGGCCC 486  
Db 135222 TGCCCTGCCACCCACCATCTGCTGTGATGGGAGGAGCCCCACCCCTACAGGGCCC 135163  
QY 487 CTGCACCTCTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAAACCGGAGTTCGGTGG 546  
Db 135162 CTGCACCTCTCAGCTACGGGACCCCTGAGCAACAGCTGGAGCTGAACCGGGAATCTGTGG 135103

QY 547 GCGACCCCAAAACAGAACCATCTTTCGACAGTGAACCTGATGATAGTCCAGGCTGGGCGG 606  
Db 135102 GCGACCCCTTAACCGGACCATCTTTCGACAGTGAACCTTATAGACGACCATGCTGGGGGG 135043  
QY 607 CCGCTGCCCCCCCCAGCAGTAACCTCGGCGCATCAGCGCCACGCTGCTACGCGCGGGCGG 666  
Db 135042 CCGCTGCTCCCCCGCAGCAGTAACCTCGGCGCATCAGCGCCACCTGCTACAGCAGCGGTGGCGG 134983  
QY 667 CATGAGGGGGCGCGGCCACCTACAGGAGGTCATCGGCCACTACCCGGGGTCTCCTT 726  
Db 134982 CATGAGGGGGCGCGGCCACCTACAGGAGGTCATCGGCCACTACCCGGGTCTCCTT 134923  
QY 727 CCAGCACCAGCAGAGCAGTGGCGCCCTCTTGTGTGAGGGGACCCGGCTCCACACAC 786  
Db 134922 CCAGCACCAGCAAGTAACGGGCCATCTCTCTGTAGAGGGGACCCGGCTCCATCTC 134863  
QY 787 ACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAACAGAAAGG 846  
Db 134862 GCACATTGCCCACTGGA-----GAACAAGGAGGAGAGAAACAGAAAGG 134818  
QY 847 ACACCTCTCTAGGTCCCGAGGGGGCGGGCTGGGCTGGTAGTGAAGAGCAG--- 904  
Db 134817 TCACCCCTCTAGGAGTGGGGCGCGGCTGTGAGCAAAACCGCAAAAAGAGAA 134758  
QY 905 -----AACACTCGCGCTTCTTAAAGAGAGTGAGAGGAGCGGGGGCGCGAG 954  
Db 134757 AGGAAAAAAACACTCGGCATCTTAAAGAGAGAGAGAGAGGAGTCAAGGAGACAC- 134699  
QY 955 CAACGCAFCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTATCATGTGATGT 1014  
Db 134698 -CAGGCTGAGTGGCGGTGTGTGTCTCTCTGTGTGTATATAATATTATCATGTCTGT 134640  
QY 1015 CTGTCTGAATGCACAGCTAAGAGAGCTTCCAAAAAAGAGAGAGAGAGAGAGAGAG 1061  
Db 134639 GTGTCTGAATGCAGAGCTCAAAAGAGCTCAAAAGAGAGAGAGAGAGAGAGAGAG 134593

RESULT 14

AC110189

LOCUS

DEFINITION

AC110189

VERSION

HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 175754)

Birren,B., Nusbaum,C. and Lander,B.

Mus musculus, clone RP23-41204

Unpublished

REFERENCE

2 (bases 1 to 175754)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,

Brown,A., Canarata,J., Campiano,A., Chang,J., Chazaro,B.,

Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,

Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,

Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Gerde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,L., Kelle,C., LaRocque,K., Lamasares,R.,

Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,

Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,

McSwan,P., McKernan,K., Melidrim,J., Meneus,L., Mihova,T.,

Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,

Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

Sever,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Zainoun, J., Vo, A., Wilson, B., Wu, X., Wyman, B., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 175754)  
Birn, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagob, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J.,  
Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C. H., O'Connor, P., O'Donnell, P., O'Neil, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, B., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

## JOURNAL

Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Oct 21, 2002 this sequence version replaced gi:20455597.  
All repeats were identified using RepeatMasker:

Smit, A. F. A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L19642

Center clone name: 412\_Q\_4

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 172464 bases at least Q40

Consensus quality: 173124 bases at least Q30

Consensus quality: 173437 bases at least Q20

Insert size: 172000; agarose-1p

Insert size: 174354; sum-of-contigs

Quality coverage: 11.2 in Q20 bases; agarose-1p

Quality coverage: 11.0 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 8772: contig of 8772 bp in length

\* 8773 9560: contig of 100 bp

\* 9561 9660: contig of 688 bp in length

\* 9661 10202: contig of 542 bp in length

\* 10303 11757: contig of 100 bp

\* 11758 11857: contig of 1455 bp in length

\* 11858 13075: contig of 100 bp

\* 13076 13175: contig of 1218 bp in length

\* 13176 15093: contig of 1918 bp in length

\* 15094 15193: gap of 100 bp

\* 15194 19311: contig of 4118 bp in length  
\* 19312 19411: gap of 100 bp  
\* 22535: contig of 3124 bp in length  
\* 22536 22635: gap of 100 bp  
\* 22636 26244: contig of 3609 bp in length  
\* 26245 26344: gap of 100 bp  
\* 26345 33067: contig of 6723 bp in length  
\* 33068 33167: gap of 100 bp  
\* 33168 49628: contig of 16461 bp in length  
\* 49629 49728: gap of 100 bp  
\* 49729 66889: contig of 17161 bp in length  
\* 66890 90869: contig of 23880 bp in length  
\* 90870 90969: gap of 100 bp  
\* 90970 122386: contig of 31417 bp in length  
\* 122387 122486: gap of 100 bp  
\* 122487 175754: contig of 53268 bp in length.

## FEATURES

## source

Location/Qualifiers

1..175754

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/clone="RP23-41204"

/clone\_lib="RPCI-23 Female Mouse BAC"

1..8772

/note="assembly\_fragment"

clone end:SP6

vector side:left

8873..9560

/note="assembly\_fragment"

9661..10202

/note="assembly\_fragment"

10303..11757

/note="assembly\_fragment"

11858..13075

/note="assembly\_fragment"

13176..15093

/note="assembly\_fragment"

15194..19311

/note="assembly\_fragment"

19412..22535

/note="assembly\_fragment"

22636..26244

/note="assembly\_fragment"

26345..33067

/note="assembly\_fragment"

33188..49628

/note="assembly\_fragment"

49729..66889

/note="assembly\_fragment"

66990..90869

/note="assembly\_fragment"

90970..122386

/note="assembly\_fragment"

122487..175754

/note="assembly\_fragment"

clone end:T7

vector side:right

BASE COUNT 42511 a 44781 c 44000 g 43050 t 1412 others

## ORIGIN

Query Match 40.8%; Score 433.4; DB 2; Length 175754;

Best Local Similarity 77.3%; Pred. No. 8.9e-72;

Matches 593; Conservative 0; Mismatches 136; Indels 38; Gaps 4;

QY 307 CCCAGAGCCGAGGCTTACGCCCGCTCGGCCACCGACCGCTGGCGCGCCCTTT 366

Db 146924 CCTCCAGCCACAGGCTCTATGCCCGCTCGGCCACTGACGACTCGCTGTGCCCTTT 146983

QY 367 CGCCCGGCGGAGCGCTTCCACCGCTTCCAGCCCGCTTCCGTTACCTGCAGCAGAT 426

Db 146984 CATCCGCGGAG-----CCGATTCCAAACCCACCTACCTACCTGCAGCAGAAAT 147034

```
Qy 427 CGACCTCGCCGCCACCATCTGCTGTGACAGCGGAGGAGCCGCCACCTTACAGGGCCC 486
Db 147035 TGCCCTGCGCCACCATCTGCTGTGATGGGAGGAGCCGCCACCTTACAGGGCCC 147094
Qy 487 CTGACCCCTTCCAGCTTGGGAGCCCGGAGCAGCAGCTGGAACCTGAACCCGGGAGTCCGGTGG 546
Db 147095 CTGACCCCTTCCAGCTTGGGAGCCCGGAGCAGCAGCTGGAACCTGAACCCGGGAGTCCGGTGG 147154
Qy 547 CGACCCGCCAACAGACCATCTTGCAGTAGTACCTGATGATGATGATGATGATGATGATGATGATG 606
Db 147155 CGACCCGCCAACAGACCATCTTGCAGTAGTACCTGATGATGATGATGATGATGATGATGATG 147214
Qy 607 CCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCCAGCAGCTGCTAGCGAGCGGGGGGG 666
Db 147215 CCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCCAGCAGCTGCTAGCGAGCGGGGGGG 147274
Qy 667 CATGAGGGGGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGGTCTCTCTT 726
Db 147275 CATGAGGGGGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGGTCTCTCTT 147334
Qy 727 CGACACGACGAGCAGTAGTGGCGCCCTCTTGTGAGGGGAGCCGGCTCCACACAC 786
Db 147335 CGACACGACGAGCAGTAGTGGCGCCCTCTTGTGAGGGGAGCCGGCTCCACACAC 147394
Qy 787 ACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGATTAACACAGAGG 846
Db 147395 GCACATTCGCCCCACTGGA-----GAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 147439
Qy 847 ACACCTCTCTTAGGTGTCGCCAGGGGGCGGGCTGGGTGCTAGGTGAAAGGAGCAG-- 904
Db 147440 TCACCCCTCTAGGTGTCGCCAGGGGGCGGGCTGGGTGCTAGGTGAAAGGAGAGAGAGAGAG 147499
Qy 905 -----ACACCTCGCGCTCTTATGAAGAGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 954
Db 147500 AGGAAAAAACAACACTCCGCACCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 147558
Qy 955 CAACGATCGTGGCCCTCCCTCCACCTCCCTGCTGTATATAATTTACATGTGATGT 1014
Db 147559 -CAGGCTGATGGCGGTGGTAGTTCTCTCTGCTGTATATAATTTACATGTGATGT 147617
Qy 1015 CTGGTCTGAATCAACAAGCTAAGAGAGCTTGAACAAAAAAGAGAGAGAGAGAGAGAGAGAG 1061
Db 147618 GTGGTCTGAATCAACAAGCTAAGAGAGCTCAAAAGCTTGAACAAAAAAGAGAGAGAGAGAG 147664
```

```
RESULT 15
AL837520/c
LOCUS 176821 bp DNA linear HTG 01-JUN-2003
DEFINITION Mus musculus chromosome 2 clone RP23-41204, *** SEQUENCING IN
PROGRESS ***
AL837520
AL837520.25 GI:31335607
VERSION HTG; HTGS PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 176821)
AUTHORS Bates,K.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 2, 2003 this sequence version replaced gi:30962480.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
```

```
----- Project Information
Center project name: BM41204
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 99% of reads
Consensus quality: 176788 bases at least Q40
Consensus quality: 176819 bases at least Q30
Consensus quality: 176821 bases at least Q20
Insert size: 176821; sum-of-ctnigs
Insert size: 180785; 10.3% error; agarose-fp
Quality coverage: 9.74x in Q20 bases; sum-of-ctnigs Quality
coverage: 9.53x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 176821: contig of 176821 bp in length.
```

## FEATURES

source

```
1..176821
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-41204"
/clone_lib="RPCI-23"
1..176821
/note="assembly_fragment:00029
vector_side:right"
```

## misc\_feature

```
1..176821
/note="assembly_fragment:00029
vector_side:right"
```

```
BASE COUNT 43349 a 44412 c 45766 g 43294 t
ORIGIN
```

```
Query Match 40.8%; Score 433.4; DB 2; Length 176821;
Best Local Similarity 77.3%; Pred. No. 8.9e-72;
Matches 593; Conservative 0; Mismatches 136; Indels 38; Gaps 4;
Qy 307 CCCAGAGCGGAGGCTTACGCGCCCGCTCGCCGCCACCGAGCGCTGGCGGCGCCCTT 366
Db 28831 CCTCAGCCAGGCTTATGCGCCCGCTCGCCGCCACCTACCGGCTGCGTGGCCCCCTT 28772
Qy 367 CGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCAGAGAT 426
Db 28771 CATCAGCGGAG-----CCGATTCACACCCACCTACCCCTACCTGCAGCAGGAAT 28721
Qy 427 CGACCTGCGCCCGCCACCATCTCGTGTGTAGAGCGGGAGAGCGCCCGCCACCTACAGGGCCC 486
Db 28720 TGCCCTGCGCCCGCCACCATCTCACTGTCTGTATGGGAGAGAGCGCCCGCCACCTACAGGGCCC 28661
Qy 487 CTGACCCCTCAGCTTGGGACCCCGCAGCAGCAGCTGGAATGACCGGGAGTCCGGTGG 546
Db 28660 CTGACCCCTCAGCTTGGGACCCCGCAGCAGCAGCTGGAATGACCGGGAGTCTGTGG 28601
Qy 547 CGCACCCCAACAGAACCATCTTCGACAGTGCACCTGTATGGATAGTGCAGGCTGGGGGG 606
Db 28600 CGCACCCCAACAGAACCATCTTCGACAGTGCACCTGTATAGACAGCAGCAGTGTGGGG 28541
Qy 607 CCCCTGCGCCCGCCAGCAGTAACCTCGGGCATCAGCGGCCAGCTGTACGGAGCGGGGGGG 666
Db 28540 CCCCTGTCGCCCGCCAGCAGTAACCTCGGGCATCAGCGGCCAGCTGTACAGCAGCGGGGGGG 28481
Qy 667 CATGAGGGGGCGCGCCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGGTCTCTT 726
Db 28480 CATGAGGGGGCGCGCCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGGTCTCTT 28421
Qy 727 CCAGCACCAGCAGCAGTGGGGCGCCCTCTCTTGTGTGAGGGGGAGCCCGGCTCCACACAC 786
```

Db 28420 CCAGCACGCAAGTAACGGGCCATCTCCCTGTAGAGGGACCCGGCTCCATCACTC 28361  
Qy 787 ACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGG 846  
Db 28360 GCACATTTGCCCCCACTGGA-----GAACAAGGAGAGAGGAAACAGAAAGG 28316  
Qy 847 ACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGCTGCGTAGGTGAAAAGGCAG-- 904  
Db 28315 TCACCCCTCTTAGGAGTGGGGGCCGGGGCGCTGTAGGCAAAACCGCAAAAAAAGAA 28256  
Qy 905 -----AACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAG 954  
Db 28255 AGGAAAAAACAACACTCCGCACTTCTTAAGAGAGAGAGAGGAAGTCAGGGGACACA- 28197  
Qy 955 CAACCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGATGT 1014  
Db 28196 -CAGGCTGAGTGGCCGTGTGGTAGTTCTCCTGTGTGTATATAATTTACATGTCTGT 28138  
Qy 1015 CTGGTCTGAATGCACAGCTAAGAGAGCTTGCAAAAAAAGAA 1061  
Db 28137 GTGGTCTGAATGCAGAGCTCAAAAGCTTGCAAAAAAAGAA 28091

Search completed: December 8, 2003, 03:10:56  
Job time : 4140 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 01:32:55 ; Search time 80 Seconds  
(Without alignments)  
5853.844 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tcttccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1022.4	96.4	1140	4	US-09-769-482-1
2	757.4	71.4	759	4	US-09-769-482-2
3	349	32.9	921	4	US-09-091-952A-7
4	349	32.9	8065	4	US-09-091-952A-6
5	281	26.5	867	4	US-09-091-952A-8
6	47.8	4.5	696	4	US-09-252-991A-13687
7	47.8	4.5	1452	4	US-09-252-991A-13650
8	47.8	4.5	9438	1	US-07-945-283-1
9	47.4	4.5	1215	4	US-09-266-965-42
10	47.4	4.5	53500	4	US-09-266-965-76
11	46.2	4.4	561	4	US-09-252-991A-4586
12	46.2	4.4	1203	4	US-09-252-991A-4391
13	46.2	4.4	1611	4	US-09-252-991A-4200
14	46.2	4.4	13778	3	US-08-785-420-1
15	46	4.3	3756	4	US-09-252-991A-692
16	46	4.3	3774	4	US-09-252-991A-719
17	46	4.3	6885	4	US-09-252-991A-660
18	45.6	4.3	1119	4	US-09-252-991A-7780
19	45.6	4.3	1887	4	US-09-252-991A-7920
20	45.6	4.3	43377	2	US-08-804-227C-7
21	45.6	4.3	43377	2	US-08-804-198-1
22	45.4	4.3	424	3	US-08-127-721A-7
23	45.4	4.3	424	3	US-08-485-246A-7
24	45.4	4.3	424	3	US-09-252-991A-6427
25	44.8	4.2	534	4	US-09-252-991A-6203
26	44.8	4.2	1272	4	US-09-252-991A-15072
27	44	4.1	1140	4	US-09-252-991A-15072

28	44	4.1	1212	4	US-09-252-991A-14551	Sequence 14551, A
29	44	4.1	1242	4	US-09-266-965-72	Sequence 72, Appl
30	44	4.1	1293	2	US-08-924-440-1	Sequence 1, Appl
31	43.8	4.1	424	2	US-08-476-176B-5	Sequence 5, Appl
32	43.8	4.1	424	2	US-08-476-176B-9	Sequence 9, Appl
33	43.8	4.1	424	3	US-08-127-721A-5	Sequence 5, Appl
34	43.8	4.1	424	3	US-08-127-721A-9	Sequence 9, Appl
35	43.8	4.1	424	3	US-08-485-246A-5	Sequence 5, Appl
36	43.8	4.1	424	3	US-08-485-246A-9	Sequence 9, Appl
37	43.8	4.1	2377	4	US-09-920-668-3	Sequence 3, Appl
38	43.8	4.1	6854	4	US-09-194-905-7	Sequence 7, Appl
39	43.6	4.1	1182	4	US-09-252-991A-565	Sequence 565, App
40	43.6	4.1	2181	4	US-09-252-991A-493	Sequence 493, App
41	43.6	4.1	2196	4	US-09-252-991A-609	Sequence 609, App
42	43.2	4.1	2793	1	US-08-203-747-1	Sequence 1, Appl
43	43.2	4.1	2793	1	US-08-458-298-1	Sequence 1, Appl
44	43.2	4.1	3705	2	US-08-474-379C-64	Sequence 64, Appl
45	43.2	4.1	3705	3	US-09-146-249A-64	Sequence 64, Appl

## ALIGNMENTS

RESULT 1  
US-09-769-482-1  
; Sequence 1, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1140  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (95)..(850)  
US-09-769-482-1

Query Match 96.4%; Score 1022.4; DB 4; Length 1140;  
Best Local Similarity 99.1%; Pred. No. 3.le-230;  
Matches 1049; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

Qy	4	TCCTTGGGTTCCGGGTGAAAGCGCTTGGGGGTTTCAGTGGGCCATGATCCCCAGCTGCTGG	63
Db	1	TCCTTGGGTTCCGGGTGAAAGCGCTTGGGGTTC--GTGGCCATGATCCCCAGCTGCTGG	58
Qy	64	AGAACTGAAGCGGACCGGTCTCTCGGAAACACAGGCAATGGCGGAGCTGGAGTTGTTCA	123
Db	59	AGAACTGAAGCGGACAGTCTCTCGGAAAC--AGGCAATGGCGGAGCTGGAGTTGTTCA	117
Qy	124	GATCATCATCATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	183
Db	118	GATCATCATCATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	177
Qy	184	CCACTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGCGGAGGAGAGA	243
Db	178	CCACTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGCGGAGGAGAGA	237
Qy	244	AGATGCCCTGTCTCTCAGAAAGGATGCTGTGTCCTCGGAGAGACAGTGTTCAGGCAACGG	303

Db 238 AGATCCCTGTCTCAGAGGATGCTGTGGCCCTCGAGAGCAGTGTGAGCAACGG 297  
Qy 304 AATCCAGAGCGCAGGTCTACGCCCGCTCGGCCACCGACCCCTGGCGCTGCCGCC 363  
Db 298 AATCCAGAGCGCAGGTCTACGCCCGCTCGGCCACCGACCCCTGGCGCTGCCGCC 357  
Qy 364 CTTCCGCCAGCGGAGCGCTTCCAGCGCTTCCAGGCCACCTATCCGTACCTGCACAGA 423  
Db 358 CTTCCGCCAGCGGAGCGCTTCCAGCGCTTCCAGGCCACCTATCCGTACCTGCACAGA 417  
Qy 424 GATCAGCTCGGCCCGCCACCATCTCGCTGTGACAGCGGAGGAGCGGCCACCCCTACACGG 483  
Db 418 GATCAGCTCGGCCCGCCACCATCTCGCTGTGACAGCGGAGGAGCGGCCACCCCTACACGG 477  
Qy 484 CCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGTGGAACTGAAACCGGAGTCCGT 543  
Db 478 CCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGTGGAACTGAAACCGGAGTCCGT 537  
Qy 544 GCGCGACCCCGCCAAACAGAACCATCTTCGACAGTCACTGATGATGATGATGATGATGATG 603  
Db 538 GCGCGACCCCGCCAAACAGAACCATCTTCGACAGTCACTGATGATGATGATGATGATGATG 597  
Qy 604 CCGCCCTGCGCCCGCCAGCAGTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGCG 663  
Db 598 CCGCCCTGCGCCCGCCAGCAGTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGCG 657  
Qy 664 GCGCATGAGGCGCGCGCCCGCCACCTACAGCAGAGTCACTCGGCCACTACCCGGGTCTCTC 723  
Db 658 GCGCATGAGGCGCGCGCGCCCGCCACCTACAGCAGAGTCACTCGGCCACTACCCGGGTCTCTC 717  
Qy 724 CTTCCAGCAGCAGCAGCAGTGGCGCCCTCTCTGCTGGAGGGAGCCCGCTCCACCA 783  
Db 718 CTTCCAGCAGCAGCAGCAGTGGCGCCCTCTCTGCTGGAGGGAGCCCGCTCCACCA 777  
Qy 784 CACACATCGCGCCCTAGAGCGCAGCAGTGGAGCAAGAGAGGATAAAGAGAA 843  
Db 778 CACACATCGCGCCCTAGAGCGCAGCAGTGGAGCAAGAGAGGATAAAGAGAA 837  
Qy 844 AGGACACCTCTCTAGGCTCCCGAGGGCGCGCGCTGGCGCTGAGTGAAGAGCA 903  
Db 838 AGGACACCTCTCTAGGCTCCCGAGGGCGCGCGCTGGCGCTGAGTGAAGAGCA 897  
Qy 904 GAACACTCGCGCTCTTAGAGAGGAGTGAAGAGGAGCGCGCGCGCGCGAGCAACGCATC 963  
Db 898 GAACACTCGCGCTCTTAGAGAGGAGTGAAGAGGAGCGCGCGCGCGCGAGCAACGCATC 957  
Qy 964 GTGTGCGCTCCCTCCACCTCCCTGTGTATATATATATATATATATATATATATATATAT 1023  
Db 958 GTGTGCGCTCCCTCCACCTCCCTGTGTATATATATATATATATATATATATATATATAT 1017  
Qy 1024 ATGCACAGCTAAGAGAGCTTCAAAAAA 1061  
Db 1018 ATGCACAGCTAAGAGAGCTTCAAAAAA 1055

RESULT 2

US-09-769-482-2  
; Sequence 2, Application US/09769482  
; Patent No. 6566130  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/09769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045

; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 759  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-769-482-2

Query Match 71.4%; Score 757.4; DB 4; Length 759;  
Best Local Similarity 99.9%; Pred. No. 2.7e-168;  
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 101 ATGCGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGGTGGTATGATGATGATGATG 160  
Db 1 ATGCGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGGTGGTATGATGATGATGATG 60  
Qy 161 GTGTGATCATCTGCTGCTGAGCCACTACAAAGTGTCTGCACGGTCTCTTCATCAGCGG 220  
Db 61 GTGTGATCATCTGCTGCTGAGCCACTACAAAGTGTCTGCACGGTCTCTTCATCAGCGG 120  
Qy 221 CACAGCCAGGCGGAGAGAGAAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCG 280  
Db 121 CACAGCCAGGCGGAGAGAGAAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCG 180  
Qy 281 GAGAGCAGTGTGAGCAACGGAATCCAGAGCGCAGAGTCTACGCCCGCTCGGCC 340  
Db 181 GAGAGCAGTGTGAGCAACGGAATCCAGAGCGCAGAGTCTACGCCCGCTCGGCC 240  
Qy 341 ACCAGCCCTGGCGCTGCGCCCTTCGCCCGAGCGGAGCGCTTCACCGCTTCAGGCC 400  
Db 241 ACCAGCCCTGGCGCTGCGCCCTTCGCCCGAGCGGAGCGCTTCACCGCTTCAGGCC 300  
Qy 401 ACCTATCGTACTGTGAGCAAGATCGAATCGAATCGAATCGAATCGAATCGAATCGA 460  
Db 301 ACCTATCGTACTGTGAGCAAGATCGAATCGAATCGAATCGAATCGAATCGAATCGA 360  
Qy 461 GAGAGCGCCCACTTACAGCGGCGCTGCGCCCTTCGCCCGAGCGGAGCGCTTCAGGCC 520  
Db 361 GAGAGCGCCCACTTACAGCGGCGCTGCGCCCTTCGCCCGAGCGGAGCGCTTCAGGCC 420  
Qy 521 CTGGAATGAAACCGGAGTCTGCGCGCAACCCCAACAGAACCATCTTCAGACAGTAC 580  
Db 421 CTGGAATGAAACCGGAGTCTGCGCGCAACCCCAACAGAACCATCTTCAGACAGTAC 480  
Qy 581 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640  
Db 481 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
Qy 641 GCCAGCTGCTACGCGAGCGGCGCATGAGGGGCGCGCCCGCCACCTACAGCGAGTTC 700  
Db 541 GCCAGCTGCTACGCGAGCGGCGCATGAGGGGCGCGCGCCCGCCACCTACAGCGAGTTC 600  
Qy 701 ATCGGCCACTACCGGGGTCTCTCTTCAGACACAGCAGAGCAGTGGGCGCGCTCTCTTG 760  
Db 601 ATCGGCCACTACCGGGGTCTCTCTTCAGACACAGCAGAGCAGTGGGCGCGCTCTCTTG 660  
Qy 761 CTGAGAGGAGCGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTG 820  
Db 661 CTGAGAGGAGCGCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTG 720  
Qy 821 AGCAAGAGAGGATAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 859  
Db 721 AGCAAGAGAGGATAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 759

RESULT 3

US-09-091-952A-7  
; Sequence 7, Application US/09091952A  
; Patent No. 6458532  
; GENERAL INFORMATION:  
; APPLICANT: Detera-Wadleigh, Sevilla D.  
; Gershon, Elliot S.

Badner, Judith A.  
Goldin, Lynn R.  
Berrettini, Wade H.  
Yoshikawa, Takeo  
Sanders, Alan R.  
Esterling, Lisa E.  
TITLE OF INVENTION: Chromosomal Markers and Diagnostic  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,952A  
FILING DATE: 19-Apr-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,278  
FILING DATE: 28-OCT-1996  
APPLICATION NUMBER: PCT/US97/19381  
FILING DATE: 28-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 015280-297100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX: <unknown>  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..921  
OTHER INFORMATION: Clone 22 coding region  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-091-952A-7

Query Match 32.9%; Score 349; DB 4; Length 921;  
Best Local Similarity 73.7%; Pred. No. 9.7e-73;  
Matches 474; Conservative 0; Mismatches 160; Indels 9; Gaps 2;  
QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTGATGATGGTGGTGG 164  
DB 182 CGGAGCTGGAGTTGCCCAATCATCATCATCGTGGTGTGATGATGGTGGTGG 241  
QY 165 TGATCAGTGTGCTGAGGCCACTACAGCTGTCTGACGGTCTTCATCAGCGGCACA 224  
DB 242 TCATCGTCTGCTGTAACCACTACAAAGTCTCCACGGGTCTTCATCAACCGCCGA 301  
QY 225 GCACGGCGGAGGAGAGAGATGCCCTCTCTCAGAGGATGCTGTGGCCCTCGGAGA 284  
DB 302 ACCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361  
QY 285 GCACAGTGTGAGCAACGGAATCCACAGAGCCGAGGTCTACGCCCGGCTCGGCCACCG 344  
DB 362 GGGCCGACCGCGGTGGG-----CGCTCGGAGATCATGATGCCCGGCTCGAGG 415  
QY 345 ACCGCTGGCGTGGCGCCCTTCGCCACGAGGAGCGCTTCCACCGCTTCCAGCCCACT 404

DB 416 ACAGGTTACAGCGCGCGTCTTCATCCAGAGGGATCGCTTCAGCGCTTCCAGCCCACT 475  
QY 405 ATCCGTACTCGAGCAGCAGATCGACCTGCGCGCGCCACCATCTCGCTGTTCAGACGGGAGG 464  
DB 476 ACCCTATGTGAGCAGCAGAGATTGATCTTCTCCACCATCTCCCTGTCCGACGGTGAAG 535  
QY 465 AGCCCCCACCCTACAGGGCCCTGACACCTCTCCAGTTCGGGACCCCGAGCAGCAGTGG 524  
DB 536 AGCCACCTCTTACCCAGGGCCCTGACACCTGACCTCGAGTCCGGACCTGAACAGCAGATGG 595  
QY 525 AACTGAACCGGAGTCTCGTGGCGGCGCACCCCAACAGAACCATCTCTCCACAGTACCTGA 584  
DB 596 AACTCAACCGAGAGTCTCGTGGGGCCCGCCACCAACCGAACCATATTTGACAGTGAATTA 655  
QY 585 TGGATAGTCCAGG---CTGGCGCGCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 641  
DB 656 TAGACATTGCTATGTATAGCGGGGTCCATGCCACCCAGCAGCAGTCTGGGCGTCACTG 715  
QY 642 CCAGTGTCTACGGCAGCGCGCGCGCATGAGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 701  
DB 716 CAAGCACCTGCAGCAGTAACCGGAGGATGGAGGGGCGCGCGCGCGCGCGCGCGCGCGCG 775  
QY 702 TCGGCCACTACCGGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 744  
DB 776 TGGGCCACCCAGCGCGCTCTTCTCTCCATCACCAGCGCAG 818

## RESULT 4

US-09-091-952A-6  
; Sequence 6, Application US/09091952A  
; Patent No. 6458532  
; GENERAL INFORMATION:  
; APPLICANT: Detera-Wadleigh, Sevilla D.  
; Gershon, Elliot S.  
; Badner, Judith A.  
; Goldin, Lynn R.  
; Yoshikawa, Takeo  
; Sanders, Alan R.  
; Esterling, Lisa E.  
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091,952A  
; FILING DATE: 19-Apr-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,278  
; FILING DATE: 28-OCT-1996  
; APPLICATION NUMBER: PCT/US97/19381  
; FILING DATE: 28-OCT-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Timothy L.  
; REGISTRATION NUMBER: 35,367  
; REFERENCE/DOCKET NUMBER: 015280-297100US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8065 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1...8065
; OTHER INFORMATION: Clone 22
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116...1036
; OTHER INFORMATION: Clone 22 coding region
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 452...505
; OTHER INFORMATION: alternatively spliced portion
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5595...5685
; OTHER INFORMATION: amplified region for genotyping
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-091-952A-6

Query Match 32.9%; Score 349; DB 4; Length 8065;
Best Local Similarity 73.7%; Pred. No. 1.9e-72;
Matches 474; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTGATGATGGTGGTGG 164
DB 297 CGGAGCTGGAGTTGCGCCAAATCATCATCGTGGTGTGATGATGGTGGTGG 356

QY 165 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGCAGCGTCTTCATCAGCCGGCACA 224
DB 357 TCATCGTCTGCTGTGAGCACTACAAGTCTCCAGCGGTCTTCATCAACGCCGCA 416

QY 225 GCCAGGGCGGAGGAGAGAGATGCTCTCTCAGAGGATGCTGTGGCCCTCGGAGA 284
DB 417 ACCAGAGCCGGAGGGAGGAGCGGGCTGCCAGAGAGGGTGCCTGTGGCTTCAGACA 476

QY 285 GCACAGTGTGAGCAACGAATCCAGAGCCGAGTCTACGCCCGCTCGGCCACCG 344
DB 477 GCGCGCACCGCGCTGGG-----CGCTCGAGATCATGATGCCCGCGCTCAAGG 530

QY 345 ACCGCTGCGCGTGGCGCCCTTCGCCAGCGGAGCGCTTCACCGCTTCCAGCCCACT 404
DB 531 ACAGGTTACAGCGCGTCTTCATCATCAGAGGATCGCTTCAGCGCTTCCAGCCCACT 590

QY 405 ATCCGTACCTGAGCAGAGATGACCTCGCCGCCACCATCTCGCTGTGACAGGGGAGG 464
DB 591 ACCCTTATGTGAGCAGCAGATGATCTTCTCCACCATCTCTCCGTCCGACGTGAAG 650

QY 465 AGCCCCACCTTACAGGSCCTTCACCTTCAGCTTCGGACCCCGAGCAGAGCTGG 524
DB 651 AGCCACCTCTTACAGGGGCGCTTCACCTTCAGCTTCGGACCCCTGAACAGAGATGG 710

QY 525 AACTGAACCGGAGTGGTGGCGGACCCCAACAGAACCATCTTCAGACGTGACCTGA 584
DB 711 AACTCAACGAGATCGGTGGGGCCCCCAACCAACCAATATTGACAGTATTAA 770

QY 585 TGATAGTGCAGG---CTGGGGGGCCCTTGGCCCCCGAGAGTAACTCGGGCATCAGG 641
DB 771 TAGACATTGATATAGCGGGGGTCCATGCCCCACCCAGCAGCAACTCGGGCATCAGT 830

QY 642 CCAGGTGCTACGAGCGCGGCGCATGAGGGGCGCGCCCACTTACAGCGAGGTCA 701
DB 831 CAAGCACTTGACAGTAAAGGAGGATGGAGGGGCCACCCCCACATACAGCGAGTGA 890

QY 702 TCGGCCACTACCGGGGGTCTCTTTCAGCACCAGCAGAGCAG 744
DB 891 TGGGCCACCACCGCGCTCTTTCTTCATCACCCAGGCGAG 933

RESULT 5
US-09-091-952A-8
; Sequence 8, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic
; Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1...867
; OTHER INFORMATION: Clone 22 isoform 2 alternatively
; spliced coding region
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-091-952A-8

Query Match 26.5%; Score 281; DB 4; Length 867;
Best Local Similarity 69.2%; Pred. No. 7.7e-57;
Matches 445; Conservative 0; Mismatches 135; Indels 63; Gaps 2;

QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGTGATGATGGTGGTGG 164
DB 182 CGGAGCTGGAGTTGCGCCAAATCATCATCGTGGTGTGATGATGGTGGTGG 241

QY 165 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGCAGCGTCTTCATCAGCCGGCACA 224
DB 242 TCATCGTCTGCTGTGAGCAACTACAAGTCTCCAGCGGTCTTCATCAACGCCCGCA 301
```

QY 225 GCCAGGGCGGAGGAGAGATGCCCTGTCTCTCAGAAAGATGCTGTGGCCCTCGGAGA 284  
DB 302 ACCAGAGCGCGGAGGAGGAGCGGCT-----  
QY 285 GCACAGTGTGACGCAACGAATCCAGAGCGCGCAGGTCTACGCCCGGCTGGGCCACCG 344  
DB 330 -----GCCGCAGATCATGATGCCCGCGGTCCAGGG 361  
QY 345 ACCGCTGTGCGCTGCGCGCTTCGCCAGCGGAGCGCTTCACCGCTTCAGACCCACCT 404  
DB 362 ACAGGTTACAGCGCGCTCTTCATCCAGAGGAGTGGTTACAGCGCTTCAGACCCACCT 421  
QY 405 ATCCGTACTGTGAGCAGAGATCGACCTGCGCCGCCACCATCTCGTGTGACAGCGGGAGG 464  
DB 422 ACCCTATGTGAGCAGCAGAGATGATCTTCCTCCACCATCTCCCTGTCCGACGGTGAAG 481  
QY 465 AGCCCCCACCCTACAGAGCGCCCTCCACCTTCAGCTTCGGACCCCGAGCAGCAGCTGG 524  
DB 482 AGCCACCTCTTACCAGGGGCGCTTCACCTCGACCTCGGAGCCCTGAACAGCAGATGG 541  
QY 525 AACTGAACCGGAGTGTGCGCGCACCCGCCAAACAGAACCATCTTCGACAGTGAACCTGA 584  
DB 542 AACTCAACGAGAGTCCGTGAGGGCCCCACCCAAACGAAACCATATTTGACAGTGAATTAA 601  
QY 585 TGGATAGTGCAGG---CTGGGCGGCGCTGCCCCCGCAGCAGTAACTCGGCGCATCAGCG 641  
DB 602 TAGACATTTGCTATGATATAGCGGGGTCCATGTCGCCACCCAGCAGCAACTCGGCGCATCAGTG 661  
QY 642 CCACGTGTACGGCAGCGCGCGCGCATGGAGGGCGCGCGCCGCCACCTACAGCAGGTCA 701  
DB 662 CAAGCACCTGACGAGTAACCGGAGGATGGAGGGGCGCCACCCGCCACATACAGCAGGTGA 721  
QY 702 TCGGCACTACCGCGGCTCTCTTCCAGCAGCAGCAGAGCAG 744  
DB 722 TGGGCCACACCCAGCGCGCTCTTCTCCATCACCAGCGCAG 764

RESULT 6  
US-09-252-991A-13687  
; Sequence 13687, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 13687  
; LENGTH: 696  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13687

Query Match 4.5%; Score 47.8; DB 4; Length 696;  
Best Local Similarity 57.8%; Pred. No. 0.025;  
Matches 85; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 109 GCTGAGTTGTTTCAGATCATCATCGTGTGGTGTGATGATGTTGTTGTTGTTGTTGAT 168  
DB 487 GCTGAATCTCTGTTTCGGCATCGTCCGCGCGCGGTGTTCTGCGCGGTGGTCTGCTGTT 546  
QY 169 CACGTGCTGTGAGCCACTACAAGCTGTCTCAGCGCTCTTCATCAGCCGCGCACAGCCA 228  
DB 547 CAAGCGCTGTGCGGAGAGAAAGAACTGGCTGAGCCTGCCATCGCCGCGCAGCGCA 606  
QY 229 GGGCGGAGGAGAGAAGATGCCCTGTC 255

DB 607 TCGCGAGGCGGATAGCCGCCCTGGC 633

RESULT 7  
US-09-252-991A-13650/C  
; Sequence 13650, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 13650  
; LENGTH: 1452  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13650

Query Match 4.5%; Score 47.8; DB 4; Length 1452;  
Best Local Similarity 57.8%; Pred. No. 0.032;  
Matches 85; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 109 GCTGAGTTGTTTCAGATCATCATCGTGTGGTGTGATGATGTTGTTGTTGTTGTTGAT 168  
DB 262 GCTGAATCTCTGTTTCGGCATCGTCCGCGCGCGGTGTTCTGCGCGGTGGTCTGCTGTT 203  
QY 169 CACGTGCTGTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGCGCACAGCCA 228  
DB 202 CAAGCGCTGTGCGGAGAAAGAAAGACTGGTGTGAGCCTGCCATCGCCGCGCAGCGCA 143  
QY 229 GGGCGGAGGAGAGAAGATGCCCTGTC 255  
DB 142 TCGCGAGGCGGATAGCCGCCCTGGC 116

RESULT 8  
US-07-945-283-1/C  
; Sequence 1, Application US/07945283  
; Patent No. 5352596  
; GENERAL INFORMATION:  
; APPLICANT: Cheung, Andrew K.  
; APPLICANT: Wesley, Ronald D.  
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants  
; TITLE OF INVENTION: Involving The EP0 and LLT Genes  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis P. Ribando  
; STREET: 1815 No. 5352596th University Street  
; CITY: Peoria  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 61604  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/07/945,283  
; FILING DATE: 19920911  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ribando, Curtis P  
; REGISTRATION NUMBER: 27976  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 309-685-4011 ext.513  
TELEFAX: 309-685-4128  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8438 base pairs  
TYPE: NUCLEIC ACID  
TOPOLOGY: double  
STRANDEDNESS: double  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Pseudorabies virus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 622..6495  
FEATURE:  
NAME/KEY: variation  
LOCATION: replace(1099, "g")  
FEATURE:  
NAME/KEY: variation  
LOCATION: replace(1267, "t")  
FEATURE:  
NAME/KEY: variation  
LOCATION: replace(1381, "c")  
FEATURE:  
NAME/KEY: variation  
LOCATION: replace(1566, "c")  
FEATURE:  
NAME/KEY: variation  
LOCATION: replace(7010, "g")  
US-07-945-283-1

Query Match 4.5%; Score 47.8; DB 1; Length 8438;  
Best Local Similarity 46.4%; Pred. No. 0.054;  
Matches 189; Conservative 0; Mismatches 217; Indels 1; Gaps 1;  
QY 348 GCCTGGCGTGGCGGCTTCGCCAGCGGGAGGCTTCACCGCTTCAGGCCACTATC 407  
Db 6181 GCCCGGGCGGGCGGCGGCTTCGGACCCCGGGCCACAGACAGCAGCGGCTGC 6122  
QY 408 CGTACCTGCAGCAGAGATCGACCTTCGCCGCCACCATCTCGCTGTTCAGACGGGAGGAGC 467  
Db 6121 CGCAGCGCGGACGCGGGGCGCGCTTCCTCC-CAGCTTCCTCCCGCGGCGGCGGCT 6063  
QY 468 CCCACCTTACAGGCGCCCTCGACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGGAC 527  
Db 6062 CGCGCGCCCGCGGCGGCGGCGGCTTCGGGAGGAGCGGCGGCTCGTCAACCATCAC 6003  
QY 528 TGAACCGGAGTCGGTGGCGGCACCCCAACAGAACCATCTTCAGCAGTGACCTGATGG 587  
Db 6002 AGCAGCGAGGCGCTTCGACGAGCGCGCGCGGCGGCGGCTCGATCTCTCGGCGCG 5943  
QY 588 ATAGTGCCAGGCTGGCGGCGGCTTCGCCGCCAGAGTAATCTCGGCGCATCAGCGCACT 647  
Db 5942 TTGGCGCGGAGCGGCGGCTTCCTCAACCCCACTCCAGCTTCAGCTCCACCAACCG 5883  
QY 648 GTTAGCGAGCGCGGCGGCGGCTTCAGGAGGCGGCGGCGGCGGCTTCAGCGAGTTCGCG 707  
Db 5882 TCGCGCTGCGAGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCTTCGCGCTTCGCG 5823  
QY 708 ACTACCGGCGGCTTCCTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 754  
Db 5822 GGGACGGCGGCGGCGGCGGCTTCGACAGACAGACAGCAGCAGCAGCAGCAGCAGCAG 5776

RESULT 9  
US-09-266-965-42/c  
Sequence 42, Application US/09266965  
Patent No. 6495348  
GENERAL INFORMATION:  
APPLICANT: Sherman, D  
APPLICANT: Mao, Y

APPLICANT: Varoglu, M  
APPLICANT: He, M  
APPLICANT: Sheldon, P  
TITLE OF INVENTION: Mitomycin biosynthetic gene cluster  
FILE REFERENCE: 600.456US1  
CURRENT APPLICATION NUMBER: US/09/266,965  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: US 08/624,447  
EARLIER FILING DATE: 1996-08-19  
EARLIER APPLICATION NUMBER: PCT/US94/11279  
EARLIER FILING DATE: 1994-10-06  
EARLIER APPLICATION NUMBER: US 08/133,963  
EARLIER FILING DATE: 1993-10-07  
NUMBER OF SEQ ID NOS: 145  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 42  
LENGTH: 1215  
TYPE: DNA  
ORGANISM: Streptomyces lavendulae  
US-09-266-965-42

Query Match 4.5%; Score 47.4; DB 4; Length 1215;  
Best Local Similarity 47.2%; Pred. No. 0.037;  
Matches 176; Conservative 0; Mismatches 196; Indels 1; Gaps 1;  
QY 441 CCATCTCGCTGTGACGCGGGAGGAGCCGCCACCTACAGGGCGGCTTCACCTCCAGC 500  
Db 775 CCAGTCGGTGTGCGGTGACGAGGAGCGCGGCTTCCTCCACCGTCTCCCGGGGC 716  
QY 501 TTGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTTCGGTGGCGGCACCCCAACA 560  
Db 715 CTGGCGGCGGCTGCGGAGGAGCGCGCTCCAGTCGGCGCTCGCCCGGCTCA 656  
QY 561 GAACATCTTCGACAGTACCTGTATGTAGTGTGCGAGCTGGGCGGCGGCTTCGCCCCCA 620  
Db 655 TCCAGTCGCGGACATCGTGTGCGCGCGCGGACCCAGCGCGGCGGCGGCGGCGGCGG 596  
QY 621 GCAGTAACCTCGGCGATCAGCCCACTGCTACGCGAGCGGGCGGCGGCGGCGGCGGCG 680  
Db 595 GGCCACATCTCAGCAGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 536  
QY 681 CGCCACCTACAGGAGGTCTATCGGCCACTACCGGGGTCTCTTCAGCAGCAGCAGCAG 740  
Db 535 GGTGACGAAACACCGGCTGCGGTGCGGAGTCCGCTGCGGATCCAGC-TCAGCCCG 477  
QY 741 GAGTGGCGGCGCTCTCTTGTGTGAGGGGACCCGGCTCCACACACACATCGCGCCCC 800  
Db 476 AACTGCGCAGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 417  
QY 801 TAGAGCGCGAGC 813  
Db 416 TGCCGGCGGCGC 404

RESULT 10  
US-09-266-965-76  
Sequence 76, Application US/09266965  
Patent No. 6495348  
GENERAL INFORMATION:  
APPLICANT: Sherman, D  
APPLICANT: Mao, Y  
APPLICANT: Varoglu, M  
APPLICANT: He, M  
APPLICANT: Sheldon, P  
TITLE OF INVENTION: Mitomycin biosynthetic gene cluster  
FILE REFERENCE: 600.456US1  
CURRENT APPLICATION NUMBER: US/09/266,965  
CURRENT FILING DATE: 1999-03-12  
EARLIER APPLICATION NUMBER: US 08/624,447  
EARLIER FILING DATE: 1996-08-19  
EARLIER APPLICATION NUMBER: PCT/US94/11279  
EARLIER FILING DATE: 1994-10-06  
EARLIER APPLICATION NUMBER: US 08/133,963







Qy 677 CCGCGCGCCACTACAGC 694  
Db 3532 CCGCGCGCCAGCAGCGC 3549

Search completed: December 8, 2003, 03:57:31  
Job time : 84 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 01:15:30 ; Search time 2637 Seconds  
(without alignments)  
9778.934 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctccctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_ges1.\*  
29: gb\_ges2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	889.4	83.8	1201	9	AL517150
C 2	878.4	82.8	1009	9	AL578575
C 3	858.8	80.9	1007	9	AL558882
C 4	844.2	79.6	951	9	AL558881

C 5	794.4	74.9	874	13	BX362396
C 6	776	73.1	945	13	BUS39219
C 7	773	72.9	967	13	BQ641849
C 8	768.4	72.4	782	12	BQ015170
C 9	724.4	68.3	729	13	BQ057571
C 10	720.2	67.9	730	12	BM677602
C 11	709.8	66.9	728	13	BQ683523
C 12	697.8	65.8	850	13	BQ602918
C 13	680.2	64.1	890	13	BQ690750
C 14	652	61.5	655	13	BQ691705
C 15	644.2	60.7	693	9	AI761441
C 16	636.2	60.0	646	13	BUS59841
C 17	635.2	59.9	964	13	BUS59860
C 18	612.4	57.4	629	13	BUT30650
C 19	609.2	57.4	618	14	CD367193
C 20	609	57.4	609	13	BQ636742
C 21	605	57.0	626	12	BM974296
C 22	599	56.5	602	14	CA431191
C 23	594.8	56.1	1046	12	BM922276
C 24	592.8	55.9	633	12	BM714472
C 25	591	55.7	651	14	CB554226
C 26	587.2	55.3	1079	11	BC023092
C 27	585	55.1	588	13	BQ624784
C 28	578	54.5	844	13	BQ686793
C 29	578	54.5	952	13	BUS7959
C 30	573	54.0	973	13	BUI69156
C 31	572.2	53.9	641	9	AW071693
C 32	569.4	53.7	938	13	BUI57842
C 33	562.2	53.0	570	13	BQ575582
C 34	562.2	53.0	916	13	BQ954555
C 35	559.2	52.7	674	9	AI972096
C 36	559	52.7	559	10	BE855409
C 37	555.8	52.4	619	9	AI742327
C 38	548.4	51.7	551	12	BM141979
C 39	534	50.3	547	12	BM676516
C 40	534	50.3	552	12	BM713900
C 41	534	50.3	563	14	CB049800
C 42	528.6	49.8	588	9	AI377498
C 43	523.8	49.4	1207	11	AK008976
C 44	515.8	48.6	634	9	AI826012
C 45	502	47.3	518	9	AI885001

#### ALIGNMENTS

RESULT 1  
AL517150/c  
LOCUS AL517150 Homo sapiens 1201 bp mRNA linear EST 09-MAY-2003  
DEFINITION CSODA008YB23 3-PRIME, mRNA sequence.

ACCESSION AL517150 GI:30492472

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 13, 2001 this sequence version replaced gi:12780643.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9945.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSODA008CAL2NP1&cluster=9945.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue Genoscope sequence ID : CS0DA08CA12NP1.  
FEATURES  
Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DA08YB23"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."  
BASE COUNT 177 a 372 c 401 g 214 t 37 others  
ORIGIN

Query Match 83.8%; Score 889.4; DB 9; Length 1201;  
Best Local Similarity 97.2%; Pred. No. 6.3e-157; Indels 1; Gaps 1;  
Matches 913; Conservative 1; Mismatches 24; Indels 1; Gaps 1;  
QY 82 TCTCTCGCAACACGAGGCTGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGT 141  
DB 939 TTTGTTCCAGAGCATGGAGATCAGCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 880  
QY 142 GGTGATGATGGTGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGC 201  
DB 879 GGTGATGATGGTGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGC 820  
QY 202 ACGTCTCTTCATCAGCCGCGCAGCCAGCGGGCGGAGAGAGATGCCCTGTCTTCAGA 261  
DB 819 ACGTCTCTTCATCAGCCGCGCAGCCAGCGGGCGGAGAGAGATGCCCTGTCTTCAGA 760  
QY 262 AGGATGCTGTGTGCTCGGAGAGCAGTGTGAGCAACGGAATCCAGAGCCGAGGT 321  
DB 759 AGGATGCTGTGTGCTCGGAGAGCAGTGTGAGCAACGGAATCCAGAGCCGCA-GT 701  
QY 322 CTAGCCCGCGCTCGGCCACCGACCGCTGGCGGTGGCGGCTTTCGCCAGCGGAGCG 381  
DB 700 CTAGCCCGCGCTCGGCCACCGACCGCTGGCGGTGGCGGCTTTCGCCAGCGGAGCG 641  
QY 382 CTTCCACCGCTTCAGCCGCGCCTATCTGCTGCTGAGCAGCAGATGACCTCGCGCCAC 441  
DB 640 CTTCCACCGCTTCAGCCGCGCCTATCTGCTGCTGAGCAGCAGATGACCTCGCGCCAC 581  
QY 442 CATCTCGCTGTGAGCGGGAGAGCCCGCCACCTACAGGGCCCTGACACCTCCAGCT 501  
DB 580 CTTCTCGCTGTGAGCGGGAGAGCCCGCCACCTACAGGGCCCTGACACCTCCAGCT 521  
QY 502 TGGGACCCGAGCAGCAGCTGGAATGAACCGGGAGTGGTGGCGGCAACCCCAACAG 561  
DB 520 TGGGACCCGAGCAGCAGCTGGAATGAACCGGGAGTGGTGGCGGCAACCCCAACAG 461  
QY 562 AACCATCTTCAGCAGTACTGATGATAGTGCAGGCTGGGCGGCGCTGCCCCCAG 621  
DB 460 AACCATCTTCAGCAGTACTGATGATAGTGCAGGCTGGGCGGCGCTGCCCCCAG 401  
QY 622 CAGTAACCTCGGCGATCAGCGCCACGCTGCTACGCGAGCGCGGCGCATGAGGGCGCGC 681  
DB 400 CAGTTACTCGGCGATCAGCGCCACGCTGCTACGCGAGCGCGGCGCATGAGGGCGCGC 341  
QY 682 GCCACCTCAGCGAGGTATCGGCCACTACCGGGGCTCTCTTCAGCAGCAGAGAG 741  
DB 340 GCCACCTCAGCGAGGTATCGGCCACTACCGGGGCTCTCTTCAGCAGCAGAGAG 281  
QY 742 CAGTGGCGCGCTCTCTTCAGCGAGGAGCCCGGCTCCACCAACACATCGCGCCCT 801  
DB 280 CAGTGGCGCGCTCTCTTCAGCGAGGAGCCCGGCTCCCGCCACACACATCGCGCCCT 221  
QY 802 AGAGAGCGCAGCCTCTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTAGGG 861  
DB 220 AGAGAGCGCAGCCTCTGGAGCAAGAGAGGATTAACAGAAAGGACACCTCTCTAGGG 161

QY 862 TCCCGAGGGCGCGGCTGGGCTGGCTAGGTGAAGAGGAGAGCAACTCCGCGCTTCTT 921  
DB 160 TCCCGAGGGCGCGGCTGGGCTGGCTAGGTGAAGAGGAGAGCAACTCCGCGCTTCTT 101  
QY 922 AGAAGAGAGGTGAGAGGAAGCGGGGCGGAGCAACGATCGTGTGGCCCTCCCTCCC 981  
DB 100 AGAAGAGAGGTGAGAGGAAGCGGGGCGGAGCAACGATCGTGTGGCCCTCCCTCCC 41  
QY 982 ACCTCCCTGTGTATAATATTATTTACATGATGATGTCTGGTC 1020  
DB 40 ACCNCCVCTGTATAATATTATTAACNWTGATGTCTGGTC 2

## RESULT 2

AL578575/c  
LOCUS  
DEFINITION AL578575 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DK001YC24 3-PRIME, mRNA sequence.  
ACCESSION AL578575  
VERSION AL578575.2 GI:31316780  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1009)  
Li'W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished  
COMMENT On Feb 16, 2001 this sequence version replaced gi:12942781.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9945.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DK001BB12NP1&cluster=9945.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DK001BB12NP1.

## FEATURES

source

1..1009  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK001YC24"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoRV  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 160 a 293 c 344 g 201 t 11 others  
ORIGIN

Query Match 82.8%; Score 878.4; DB 9; Length 1009;  
Best Local Similarity 96.3%; Pred. No. 7.5e-155;  
Matches 896; Conservative 11; Mismatches 22; Indels 1; Gaps 1;

QY 82 TCTCTCGCAACACGAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGT 141  
DB 929 TTTGTTCCAGAGCATGGAGATCAGCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 870  
QY 142 GGTGATGATGGTGGTGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGC 201  
DB 869 GGTGATGATGGTGGTGGTGGTGTATCATCGTCCCTGCTGAGCCACTACAAGCTGTCTGC 810  
QY 202 ACGTCTCTTCATCAGCCGCGCAGCCAGGGCGGAGGAGAGATGCCCTGTCTTCAGA 261



```

661 GCTGAGGGACCGGCTCCACCACACACATCGCGCCCTAGAGGGCA-SCATCTG 719
      |||
820 GAGCAAGAGAGAGATAAACAGAAAGACACCCCTCTAGGGTCCCAAGGGGGCGGGC 879
      |||
720 GAGCAAGAGAGAGATAAACAGAAAGACACCCCTCTAGGGTCCCAAGGGGGCGGGC 779
      |||
880 TGGGGCTCGTAGGTGAAGAGCAGACACATCCGGCTTCTTAGAAGAGGAGTGAGAGGA 939
      |||
780 TGGGGCTCGTAGGTGAAGAGCAGACACATCCGGCTTCTTAGAAGAGGAGTGAGAGGA 839
      |||
940 AGGGGGGGGGCGCAGACGATCGTGTGGCCCTCCCTCCACATCCCTCTGTGTATAAAT 999
      |||
840 AGGGGGGGGGGSM--SAACGCATCGTGTGGCCCTCCCTCCACATCCCTCTGTGTATAAAT 897
      |||
1000 ATTACATGTGATGTCTGGTGTGAATGCAAGAGTGAAGAGAGCTTGAAGAGAGAGAGAGAG 1059
      |||
898 ATTACATGTGATGTCTGGTGTGAATGCAAGAGTGAAGAGAGCTTGAAGAGAGAGAGAGAG 957
      |||
1060 AA 1061
      ||
958 AA 959

```

```

RESULT 4
AL558881/c
LOCUS
DEFINITION
AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ015YF12 3-PRIME, mRNA sequence.
ACCESSION
AL558881
VERSION
AL558881.2 GI:31283014
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12903836.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DC06NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015DC06NP1.
Location/Qualifiers
1. 951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YF12"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
141 a 275 c 318 g 201 t 16 others
ORIGIN

```

FEATURES  
source

```

160 GGTGTGATCAGCTGCTGCTGAGCCACTCAAGCTGTCTGACGGTCTCTTTCATCAGCG 219
      |||
854 GGATGTGTCACGTGCTGCTGAGCCACTCAAGCTGTCTGACGGTCTCTTTCATCAGCG 795
      |||
220 GCACAGCCAGGGGGAGGAGAGATGCGCTCTCTCAGAAAGATGCTGTGCGCCCTC 279
      |||
794 GAAACAGCCAGGGGGAGGAGAGATGCGCTCTCTCAGAAAGATGCTGTGCGCCCTC 735
      |||
280 GGAGAGACATGTGTACGCAACCGAATCCAGAGCGGAGGTCTACGCCCGCTCGGCC 339
      |||
734 GGAGAGACATGTGTACGCAACCGAATCCAGAGCGGAGGTCTACGCCCGCTCGGCC 675
      |||
340 CACGACCGCTGCGCTGCGCCCTCTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCC 399
      |||
674 CACGACCGCTGCGCTGCGCCCTCTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCC 615
      |||
400 CACCTATCCGTACCTGACAGACGAGATCGACTCGCGCCACCATCTCGCTGTTCAGACGG 459
      |||
614 CACCTATCCGTACCTGACAGACGAGATCGACTCGCGCCACCATCTCGCTTTCAGACGG 555
      |||
460 GGAGAGCGCCCACTTACAGGGCCCTGACCGCTTCCAGCTTGGGAGCCCGGAGCAGA 519
      |||
554 GGAGAGCGCCCACTTACAGGGCCCTGACCGCTTCCAGCTTGGGAGCCCGGAGCAGA 495
      |||
520 GCTGGAACCTGAACCGGGAGTGGTGGCGCACACCCCAACAGAACCATCTTCGACAGTGA 579
      |||
494 GCTGGAACCTGAACCGGGAGTGGTGGCGCACACCCCAACAGAACCATCTTCGACAGTGA 435
      |||
580 CTTGATGATAGTGCACAGGCTGGGGCGGCTTGGCGCCCTGCGCCAGCAGTAACCTCGGGCATCAG 639
      |||
434 CTTGATGATAGTGCACAGGCTGGGGCGGCTTGGCGCCCTGCGCCAGCAGTAACCTCGGGCATCAG 375
      |||
640 CGCCACGTGTACGGCAGCGCGCGCATGGAGGGCGCGCGCCCACTTACAGCGAGGT 699
      |||
374 CGCCACGTGTACGGCAGCGCGCGCATGGAGGGCGCGCGCCCACTTACAGCGAGGT 315
      |||
700 CATGGCCACTTACCGGGGTCTCTTCCAGCACCGACAGAGAGTGGGGCGCCCTCTT 759
      |||
314 CATGGCCACTTACCGGGGTCTCTTCCAGCACCGACAGAGAGTGGGGCGCCCTCTT 255
      |||
760 GCTGAGGGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTG 819
      |||
254 GCTGAGGGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCATCTG 195
      |||
820 GAGCAAGAGAGAGATAAACAGAAAGACACCCCTCTTAGGGTCCCAAGGGGGCGGGC 879
      |||
194 GAGCAAGAGAGAGATAAACAGAAAGACACCCCTCTTAGGGTCCCAAGGGGGCGGGC 135
      |||
880 TGGGGCTCGTAGGTGAAGAGCAGACACATCCGGCTTCTTAGAAGAGGAGTGAGAGGA 939
      |||
134 TGGGGCTCGTAGGTGAAGAGCAGACACATCCGGCTTCTTAGAAGAGGAGTGAGAGGA 75
      |||
940 AGGGGGGGGGCGCAGCAACGATCGTGTGGCCCTCTCCCTCCACATCCCTCTGTGTATAAAT 999
      |||
74 AGGGGGGGGGCGCAGCAACGATCGTGTGGCCCTCTCCCTCCACATCCCTCTGTGTATAAAT 15
      |||
1000 ATTACATGTGATG 1013
      |||
14 ATTTACATGTGATG 1

```

RESULT 5  
BX362396/c

```

LOCUS
DEFINITION
BX362396 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ014YN15 3-PRIME, mRNA sequence.
ACCESSION
BX362396
VERSION
BX362396.1 GI:30378625
KEYWORDS
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Query Match 79.6%; Score 844.2; DB 9; Length 951;  
Best Local Similarity 99.1%; Pred. No. 1.9e-148;  
Matches 846; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 874)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segret@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 9945.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DJ014CG08NP1&cluster=9945.r. Contact :  
Feng Liang Email: fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DJ014CG08NP1.  
Location/Qualifiers  
1. 874  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ014YN15"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 125 a 257 c 278 g 180 t 34 others  
ORIGIN

Query Match 74.9%; Score 794.4; DB 13; Length 874;  
Best Local Similarity 94.3%; Pred. No. 4.3e-139;  
Matches 824; Conservative 20; Mismatches 27; Indels 3; Gaps 3;

QY 146 ATGATGGTGTGGTGGTGTATC-ACGTGCTGTGTGAGCCACTACAAGCTGTCTGCACG 204  
DB 874 ATGATGGTGTGGTGGTGTATC-ACGTGCTGTGTGAGCCACTACAAGCTGTCTGCACG 815  
QY 205 GTCTTTCATACCGCGGACGACGAGGCGGAGGAGAGAGATGCCCTGTCTCAGAGG 264  
DB 814 GTCTTTCATACCGCGGACGACGAGGCGGAGGAGAGAGATGCCCTGTCTCAGAGG 755  
QY 265 ATGCTGTGGCCCTCGGAGAGACAGTGTTCAGGCAACGAAATCCAGAGCGGCGAGTCTA 324  
DB 754 ATGCTGTGGCCCTCGGAGAGACAGTGTTCAGGCAACGAAATCCAGAGCGGCGAGTCTA 696  
QY 325 CGCCCGGCTCGGCGGACGCGGCTGTGGCGGTGCGGCGGCTTTCGCGGCGGAGCGCTT 384  
DB 695 CGCCCGGCTCGGCGGACGCGGCTGTGGCGGTGCGGCGGCTTTCGCGGCGGAGCGCTT 636  
QY 385 CCACCGCTTCAGCGGCGGCTTTCGCTACCTGTGAGCAGCAGATCGACCTGCGCGGCCACCAT 444  
DB 635 CCACCGCTTCAGCGGCGGCTTTCGCTACCTGTGAGCAGCAGATCGACCTGCGCGGCCACCAT 576  
QY 445 CTCGCTGTTCAGCGGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 504  
DB 575 CTCGCTGTTCAGCGGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 516  
QY 505 GGACCGCGGAGCAGCTGGAACCGGAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 564  
DB 515 GGACCGCGGAGCAGCTGGAACCGGAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 456  
QY 565 CATCTTCAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 624  
DB 455 CATCTTCAGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 396  
QY 625 TAATCTGGGCGATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 684  
DB 395 TAATCTGGGCGATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 336

QY 685 CACCTACAGCGAGGTATCGGCCACTACCCGGGGTCTCTTCCAGACCCAGCAGAGCAG 744  
DB 335 CACCTACAGCGAGGTATCGGCCACTACCCGGGGTCTCTTCCAGACCCAGCAGAGCAG 276  
QY 745 TGGGCGGCGGCTCTTGTCTGGAGGGGACCCGGCTCCACACACACACACACACACAC 804  
DB 275 TGGGCGGCGGCTCTTGTCTGGAGGGGACCCGGCTCCACACACACACACACACACAC 216  
QY 805 GAGCGCAGCCATCTGGAGCAAGAGAGATATAAAGAGAGAGAGAGAGAGAGAGAGAG 864  
DB 215 GAGCGCAGCCATCTGGAGCAAGAGAGATATAAAGAGAGAGAGAGAGAGAGAGAGAG 156  
QY 865 CCAGGGGGCGGCGGCTGGGGTGTGGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 924  
DB 155 CCAGGGGGCGGCGGCTGGGGTGTGGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 96  
QY 925 AGAGGAGTGAG 984  
DB 95 AGAGGAGTGAG 36  
QY 985 -TCCCTGTGTATATAATATTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017  
DB 35 GGNNTNNTATATAATATAATAATGTGAAGAVDR 2

RESULT 6  
BUS39219  
LOCUS  
DEFINITION  
IMAGE:6569922 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BUS39219 945 bp mRNA linear EST 13-SEP-2002  
AGENCOURT 10215265 NIH\_MGC\_107 Homo sapiens cDNA clone  
IMAGE:6569922 5', mRNA sequence.  
BUS39219  
BUS39219.1 GI:22849660  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2757 row: p column: 18  
High quality sequence stop: 663.  
Location/Qualifiers  
1. 945  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6569922"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_107"  
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;  
Site 2: XhoI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 194 a 319 c 287 g 144 t 1 others  
ORIGIN

```
Query Match      73.1%; Score 776; DB 13; Length 945;
Best Local Similarity 95.7%; Pred. No. 1.2e-135;
Matches 841; Conservative 0; Mismatches 31; Indels 7; Gaps 4;

Qy 18 TGAAGCGCTTGGGGTTTCACTGGGCCATGATCCCGAGCTGCTGGAGAACTGAAGCGG 77
Db 1 TGAAGCGCTTGGGGTTTCACTGGGCCATGATCCCGAGCTGCTGGAGAACTGAAGCGG 60

Qy 78 AGGTCTCTCTGCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCG 137
Db 61 ACAGTCTCTCTGCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCG 120

Qy 138 TGGTGGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 197
Db 121 TGGTGGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180

Qy 198 CTGACGGTCTCTTATCATCGCGCACAGCCAGGGCGGAGGAGAGAAATGCTGTCT 257
Db 181 CTGACGGTCTCTTATCATCGCGCACAGCCAGGGCGGAGGAGAGAAATGCTGTCT 240

Qy 258 CAGAAGGATGCTGTGGCCCTCGGAGAGCAGATGTCTAGGCAACGGAATCCAGAGCGC 317
Db 241 CAGAAGGATGCTGTGGCCCTCGGAGAGCAGATGTCTAGGCAACGGAATCCAGAGCGC 300

Qy 318 AGGTCTACGCGCCCGCTCGGGCCACCGACCGCTGCGCGTGGCGCCCTTGGCCAGCGG 377
Db 301 AGGTCTACGCGCCCGCTCGGGCCACCGACCGCTGCGCGTGGCGCCCTTGGCCAGCGG 360

Qy 378 AGCGTCTCCAGCGCTTCCAGCGCCACCTATTCGTACTCTGAGCAGCAGATCGCGCG 437
Db 361 AGCGTCTCCAGCGCTTCCAGCGCCACCTATTCGTACTCTGAGCAGCAGATCGCGCG 420

Qy 438 CCACATCTCTGTCTAGAGCGGGAGAGCCCCACCTTACAGAGCGCGCTGACCTCC 497
Db 421 CCACATCTCTGTCTAGAGCGGGAGAGCCCCACCTTACAGAGCGCGCTGACCTCC 480

Qy 498 AGCTTCGGGACCCCGAGCAGCAGCTGGAATCTGAACCGGAGTCGTGGCGCACCCCAA 557
Db 481 AGCTTCGGGACCCCGAGCAGCAGCTGGAATCTGAACCGGAGTCGTGGCGCACCCCAA 540

Qy 558 ACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGAGTGGCGCGCCCTCGCC 617
Db 541 ACAGAACCATCTTCGACAGTACCTGATGATAGTCCAGGAGTGGCGCGCCCTCGCC 600

Qy 618 CCAGCAGTAACTCGGGCATCAGCGCCACGTGTCTAGCGAGCGGGCGCATGGAGGGC 677
Db 601 CCAGCAGTAACTCGGGCATCAGCGCCACGTGTCTAGCGAGCGGGCGCATGGAGGGC 660

Qy 678 CGCGCCACCATACAGCAGGTCATCGCCACTACCGGGTCTCTTCCAGCACCAGC 737
Db 661 CGCGCCACCATACAGCAGGTCATCGCCACTACCGGGTCTCTTCCAGCACCAGC 720

Qy 738 AGAGCAGTGGCGCGCCCT-CTTGTGTGGAGGGAGCCCGGCTCCACACACATCGCG 796
Db 721 AGAGCAGTGGCGCGCCCTCTTGTGTGGAGGGAGCCCGGCTCCACACACATCGTG 780

Qy 797 -CCCTAGAGCGGAG-CCATCTGGAGCAAGAGAGATATAA-----CAGAAAGAC 850
Db 781 CCCCCCTAAGAGCGAGCCCATCTGGAGCAAGAGAGATATAAACCCTGAAAGGAGCAC 840

Qy 851 CCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGGTGG 889
Db 841 CTCTCTTCTAGGTCCCCCAGGGGGGGGGCGCGGCTGGG 879
```

```
RESULT 7
LOCUS      B0641849
DEFINITION AGENCOURT_8287174 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6292265
           5', mRNA sequence.
ACCESSION  B0641849
VERSION    B0641849.1 GI:21766021
KEYWORDS   EST
```

```
SOURCE      Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 967)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
              Cloning by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM2493 row: g column: 18
              High quality sequence stop: 571.
```

## FEATURES

```
source
Location/Qualifiers
1..967
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6292265"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
BASE COUNT  194 a 334 c 296 g 143 t
```

```
Query Match      72.9%; Score 773; DB 13; Length 967;
Best Local Similarity 95.0%; Pred. No. 4.3e-135;
Matches 810; Conservative 0; Mismatches 40; Indels 3; Gaps 1;
```

```
Qy 107 GAGCTGGAGTTTGTTCAGATCATCATCTGCTGGTGGTGGTGGTGGTGGTGGTGG 166
Db 1 GAGCTGGAGTTTGTTCAGATCATCATCTGCTGGTGGTGGTGGTGGTGGTGGTGG 60

Qy 167 ATCAGTGCCTGCTGAGCAGCTTACAAGCTGTCTGACCGTCTTTCATCAGCGGCACAGC 226
Db 61 ATCAGTGCCTGCTGAGCAGCTTACAAGCTGTCTGACCGTCTTTCATCAGCGGCACAGC 120

Qy 227 CAGGGCGGAGGAGAGAGATGCTGTCTCAGAGGATGCTGTGGCCCTCGAGAGC 286
Db 121 CAGGGCGGAGGAGAGAGATGCTGTCTCAGAGGATGCTGTGGCCCTCGAGAGC 180

Qy 287 ACAGTGTCAAGCAACGGATCCAGAGCGGAGGCTTACGCCCGCCCTCGGCCACCCAGC 346
Db 181 ACAGTGTCAAGCAACGGATCCAGAGCGGAGGCTTACGCCCGCCCTCGGCCACCCAGC 240

Qy 347 CGCTTGGCGCGTGGCGCCCTTCCGCGGAGGAGGCTTCCACCGCTTCCAGCGCCACCTAT 406
Db 241 CGCTTGGCGCGTGGCGCCCTTCCGCGGAGGAGGCTTCCACCGCTTCCAGCGCCACCTAT 300

Qy 407 CCGTACCTGACGACAGATCGACTGCGCGCCACCATCTCGCTGTCTAGACGGGAGGAG 466
Db 301 CCGTACCTGACGACAGATCGACTGCGCGCCACCATCTCGCTGTCTAGACGGGAGGAG 360

Qy 467 CCCCACCTACAGGGGCGCTTGCACCTCTCGAGCTTCGGGACCCCGAGCAGCAGCTGGA 526
Db 361 CCCCACCTACAGGGGCGCTTGCACCTCTCGAGCTTCGGGACCCCGAGCAGCAGCTGGA 420

Qy 527 CTGAACCGGGAGTGGTGGCGGCGACCCCAACAGAACCATCTTTCAGACAGTGCATGATG 586
Db 527 CTGAACCGGGAGTGGTGGCGGCGACCCCAACAGAACCATCTTTCAGACAGTGCATGATG 586
```



Db 421 CTGAACGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATG 480  
QY 587 GATAGTGCAGAGCTGGGGGCCCCCTTGGCCCCCAGAGTAATCTGGGGCATCAGCCACG 646  
Db 481 GATAGTGCAGAGCTGGGGGCCCCCTTGGCCCCCAGAGTAATCTGGGGCATCAGCCACG 540  
QY 647 TGCTACGACGAGGGGGCGGATGAGGGGGCGCGCCACCTACAGGAGGTCTATCGGC 706  
Db 541 TGCTACGACGAGGGGGCGGATGAGGGGGCGCGCCACCTACAGGAGGTCTATCGGC 600  
QY 707 CACTACCGGGGTCTCTCTTCAGACACAGAGAGAGTGGGCCCGCCCTCTCTTGTGGAG 766  
Db 601 CACTACCGGGGTCTCTCTTCAGACACAGAGAGAGTGGGCCCGCCCTCTCTTGTGGAG 660  
QY 767 GGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGGAGCAA 826  
Db 661 GGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCCATCTGGAGCAA 720  
QY 827 GAGAGGATTAACAGAAAGGACACCTCTCTAGGTGTCCTAGGGGCGGGGCTGGG-- 884  
Db 721 GAGAGGATTAACAGAAAGGACACCTCTCTAGGTGTCCTAGGGGCGGGGCTGGG 780  
QY 885 -CTGCGTAGGTGAAAAGGAGAGACACTCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGC 943  
Db 781 TTGCTTAGGGAAAGGACACACCTCTCGCGCCCTCTTAAAGAGAGGTGGAAACGC 840  
QY 944 GGGGGGCGAGCA 956  
Db 841 AGGCGGGGGCA 853

## RESULT 8

BQ015170/c  
LOCUS  
DEFINITION BQ015170 782 bp mRNA linear EST 26-MAR-2002  
IMAGE:5834635 3', mRNA sequence.  
ACCESSION BQ015170  
VERSION BQ015170.1 GI:19740071  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 782)  
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Seq primer: M13 FORWARD  
POLYA=Yes.

## FEATURES

## source

1..782  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5834635"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP ED1"  
/note="Organ: Left Pubic Bone; Vector: p7T3-Pac  
(Pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library  
containing the following tissue(s): Chondrosarcoma cell  
line CSS. The library was constructed according to Bonaldo

Lennon and Soares, Genome Research, 6:791-806, 1996.  
First strand cDNA synthesis was primed with an oligo-dT  
primer containing a Not I site. Double stranded cDNA was  
ligated to an EcoR I adaptor, digested with Not I, and  
cloned directionally into p7T3-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GCTCAAGGCT.

TAG LIB=UI-H-ED1

TAG TISSUE=chondrosarcoma

TAG\_SEQ=CGTCAAGGCT

BASE COUNT 109 a 223 c 271 g 176 t 3 others

## ORIGIN

Query Match 72.4%; Score 768.4; DB 12; Length 782;  
Best Local Similarity 98.8%; Pred. No. 3.2e-134;  
Matches 772; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 281 GAGAGCAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 340  
Db 782 GAGAGCAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 723  
QY 341 ACCGACCGCTGGCGGCTGCGCCCTTCGCCAGCGGGAGCGCTTCCAGGCTTCAGCCC 400  
Db 722 ACCGACCGCTGGCGGCTGCGCCCTTCGCCAGCGGGAGCGCTTCCAGGCTTCAGCCC 663  
QY 401 ACCTATCCGTACCTGACGACGATCGACCTGCCGCCACCATCTCGCTGTGACAGCGG 460  
Db 662 ACCTATCCGTACCTGACGACGATCGACCTGCCGCCACCATCTCGCTGTGACAGCGG 603  
QY 461 GAGAGCCCCCACCCTACAGGGGCCCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 520  
Db 602 GAGAGCCCCCACCCTACAGGGGCCCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 543  
QY 521 CTGGAACTGAACCGGGAGTCGGTGGCGGCACACCCCAACAGAACCATCTTCGACAGTGAC 580  
Db 542 CTGGAACTGAACCGGGAGTCGGTGGCGGCACACCCCAACAGAACCATCTTCGACAGTGAC 483  
QY 581 CTGATGATAGTGTGCGGGGCGCCCTGCGCCCCCGCCAGCAGTAATCTCGGGCATCAGC 640  
Db 482 CTGATGATAGTGTGCGGGGCGCCCTGCGCCCCCGCCAGCAGTAATCTCGGGCATCAGC 423  
QY 641 GCCACGTGCTACGCGAGCGGGCGCATGAGAGGGGCGCCGCCACCTACAGCGAGGTC 700  
Db 422 GCCACGTGCTACGCGAGCGGGCGCATGAGAGGGGCGCCGCCACCTACAGCGAGGTC 363  
QY 701 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGAGTGGGGCGCCCTCCTTG 760  
Db 362 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGAGTGGGGCGCCCTCCTTG 303  
QY 761 CTGGAGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 820  
Db 302 CTGGAGGGGACCCGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 243  
QY 821 AGCAAGAGAGAGGATAAAGAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT 880  
Db 242 AGCAAGAGAGAGGATAAAGAGAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCT 183  
QY 881 GGGGCTGCTAGGTGAAAGGAGCAACACTCGCGCTTCTTAGAGAGAGAGTGTGAGAGAA 940  
Db 182 GGGGCTGCTAGGTGAAAGGAGCAACACTCGCGCTTCTTAGAGAGAGAGTGTGAGAGAA 123  
QY 941 GGGGGGGGCGCAGCAGCATCGTGGGCGCTCCCTCCACCTCCCTCTGTATATAATA 1000  
Db 122 GGGGGGGGCGCAGCAGCATCGTGGGCGCTCCCTCCACCTCCCTCTGTGTATAATA 63  
QY 1001 TTTACATGTGATGCTGTGCTCAATGCAAGAGAGAGGCTTGCACAAAAA 1060  
Db 62 TTTACATGTGATGCTGTGCTCAATGCAAGAGAGGCTTGCACAAAAA 3  
QY 1061 A 1061



cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com).  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1..730  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="UI-E-E01-aid-p-09-0-UI"  
/tissue\_type="fetal eye"  
/dev\_stages="fetal"  
/lab\_host="DHI0B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-E01"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-E01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_LIB=UI-E-E01  
TAG\_TISSUE=human fetal eye  
TAG\_SEQ=CGCGTATACC"  
BASE COUNT 103 a 210 c 251 g 164 t 2 others  
ORIGIN  
Query Match 67.9%; Score 720.2; DB 12; Length 730;  
Best Local Similarity 99.3%; Pred. No. 3.5e-125;  
Matches 722; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 335 CGGCCACCGACCGCTGGCGTGGCCGCTTTCGCCACGCGGAGCGCTTCACCGCTTC 394  
Db 730 CGGCCACCGACCGCTGGCGTGGCCGCTTTCGCCACGCGGAGCGCTTCACCGCTTC 671  
Qy 395 CAGCCCACTATCGTACTCGACGACGAGATCGACCTGCCGCCACCATCTCGCTGCA 454  
Db 670 CAGCCCACTATCGTACTCGACGACGAGATCGACCTTGCCACCCACCATCTCGCTGCA 611  
Qy 455 GACGGGAGGAGCCGCCACCCCTACGAGGCGCCCTGCACCTTCAGCTTCGGGACCCCGAG 514  
Db 610 GACGGGAGGAGCCGCCACCCCTACGAGGCGCCCTGCACCTTCAGCTTCGGGACCCCGAG 551  
Qy 515 CAGCAGCTGGAACCTGAAACGGGAGTGGTGGCGGACCCGCCAAACAGAACCATCTTCGAC 574  
Db 550 CAGCAGCTGGAACCTGAAACGGGAGTGGTGGCGGACCCGCCAAACAGAACCATCTTCGAC 491  
Qy 575 AGTGACCTGATGGATAGTGCACGCTGGCGGCGCCCTGCCCGCCCGCCAGCAGTAACCTCGGC 634  
Db 490 AGTGACCTGATGGATAGTGCACGCTGGCGGCGCCCTGCCCGCCCGCCAGCAGTAACCTCGGC 431  
Qy 635 ATCAGCGCACGCTGTACGCGAGCGGCGCGCATGGAGGGGCGCCGCCACCTACAGC 694  
Db 430 ATCAGCGCACGCTGTACGCGAGCGGCGCGCATGGAGGGGCGCCGCCACCTACAGC 371  
Qy 695 GAGGTATCGGCGCACTACCCGGGGTCTCTCTTCCAGCACCGAGCAGTGGGCGGCC 754  
Db 370 GAGGTATCGGCGCACTACCCGGGGTCTCTCTTCCAGCACCGAGCAGTGGGCGGCC 311  
Qy 755 TCCTTGTCTGGAGGGAGCCGGGTCCACACACACACATCGCGCCCTTAGAGCGCAGCC 814  
Db 310 TCCTTGTCTGGAGGGAGCCGGGTCCACACACACACATCGCGCCCTTAGAGCGCAGCC 251

Qy 815 ATCTGGAGCAAGAGAGGATAAAACAGAAAGCAGACCCCTCTCTAGGGTCCCAGGGGGC 874  
Db 250 ATCTGGAGCAAGAGAGGATAAAACAGAAAGCAGACCCCTCTCTAGGGTCCCAGGGGGC 191  
Qy 875 CGGCTGGGCTGGTAGGTGAAAAGGCGAGCACTCCGGCTTCTTTAGAGAGAGTGA 934  
Db 190 CGGCTGGGCTGGTAGGTGAAAAGGCGAGCACTCCGGCTTCTTTAGAGAGAGTGA 131  
Qy 935 GAGGAAGCGGGGGCGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 994  
Db 130 GAGGAAGCGGGGGCGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 71  
Qy 995 TAAATATTACATGTGATGTCTGTGTAATGCAACAGCTAAGAGAGCTTGCAAAAAA 1054  
Db 70 TAAATATTACATGTGATGTCTGTGTAATGCAACAGCTAAGAGAGCTTGCAAAAAA 11  
Qy 1055 AAAAAA 1061  
Db 10 AAAAAA 4  
RESULT 11  
BU683523/c  
LOCUS BU683523  
DEFINITION UI-CF-E01-acg-e-09-0-UI.s1 UI-CF-E01 Homo sapiens cDNA clone  
ACCESSION BU683523  
VERSION BU683523.1 GI:23535533  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 728)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
889548  
Contact: McCray, PB  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com) or from Open Biosystems  
(www.openbiosystems.com).  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1..728  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="UI-CF-E01-acg-e-09-0-UI"  
/tissue\_type="Lung"  
/dev\_stages="Adult and Fetal"  
/lab\_host="DHI0B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-E01"  
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-E01 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares,



```
Db 181 CTACGCCCGCTCGGCCACCGACCGCTGCGCTGCGCGCTTGGCCCGGAGCG 240
Qy 382 CTTCCACCGCTTCCAGCCACCTATCCGTAACCTGAGCAGCAGATCGACCTGCCCGCCAC 441
Db 241 CTTCCACCGCTTCCAGCCACCTATCCGTAACCTGAGCAGCAGATCGACCTGCCCGCCAC 300
Qy 442 CATCTCGCTGTCAGAGCGGAGGAGCCCGCCACCTACCGAGCGCCCTGACCTCCAGCT 501
Db 301 CATCTCGCTGTCAGAGCGGAGGAGCCCGCCACCTACCGAGCGCCCTGACCTCCAGCT 360
Qy 502 TCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTCCGTCGCGGACACCCCGCCAG 561
Db 361 TCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTCCGTCGCGGACACCCCGCCAG 420
Qy 562 AACCATCTTCGACAGTACCTGATGATAGTCCAGGCTGGCGCCCTGCCCCCGAG 621
Db 421 AACCATCTTCGACAGTACCTGATGATAGTCCAGGCTGGCGCCCTGCCCCCGAG 480
Qy 622 CAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGCGGCGCATGGAGGCGCGC 681
Db 481 CAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGCGGCGCATGGAGGCGCGC 540
Qy 682 GCCCACTTACAGCAGGCTCATCGGCACCTACCGCGGCTCCTCTTCCAGCACCAGCAG 741
Db 541 GCCCACTTACAGCAGGCTCATCGGCACCTACCGCGGCTCCTCTTCCAGCACCAGCAG 600
Qy 742 CAGTGGGCGC-CCCTCTTGTCTGG-AGGGGACCCCGCTCCACACACACATCGCGCC 799
Db 601 CAGTGGGCGCCTCTCTTGTCTGGAGGGGACCCGCTTCCACACACACATCGCGCC 660
Qy 800 CTAGAGCGCGAGCA-TCTGGAGCAAGAGAGGATTAACAGAAAGACACCTCTCTA 858
Db 661 CTAAGAGCGCGACCATCTCTGGAGCAAGCAGAAATGATAACGAAAGACCCCTCTCTA 720
Qy 859 GGGTCCCGAGGGGGCC--GGGCTGGGCTGGTAGGTGAAAGGACAGAACT---CG 913
Db 721 GGGTCCCGAGGGGGCCCGGGCTGGGCTGTCTAGGTGAAAGAGAAACATCTCCCG 780
Qy 914 CGCTCTTTAGAGAGAGTGAAGAGGCGGGGGCGGAGCAACGCTCTGTGGCCCT 973
Db 781 CTTCTTTAAAGAACGATGATAAGAAAGCGGGGGGCGCAACACCCCTTTGGGTGGCC 840
Qy 974 CCCCCTC 980
Db 841 CTCCTCC 847

RESULT 13
BQ690750
LOCUS AGENCOURT_8046394_NIH_MGC_110_Homo_sapiens_cDNA clone IMAGE:6209341
DEFINITION 5', mRNA sequence.
ACCESSION BQ690750
VERSION BQ690750.1 GI:21816066
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCM2368 row: h column: 14
```

```
High quality sequence stop: 627.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 168 a 303 c 273 g 146 t
ORIGIN
Query Match 64.1%; Score 680.2; DB 13; Length 890;
Best Local Similarity 95.9%; Pred No. 1.1e-117;
Matches 709; Conservative 0; Mismatches 28; Indels 2; Gaps 1;
Qy 1 TCCTCTTGGGTTGGGTGAAAGCGCTTGGGGGTTTCAGTGGGCCCATGATCCCGAGCTGC 60
Db 24 TCCTCTTGGGTTGGGTGAAAGCGCTTGGGGTTCAGTGGGCCCATGATCCCGAGCTGC 83
Qy 61 TGGAGAACTGAAGCGGACCGCTCTCTCGGAAACAGCAATGGCGGAGCTGGAGTTGT 120
Db 84 TGGAGAACTGAAGCGGACCGCTCTCTCGGAAACAGCAATGGCGGAGCTGGAGTTGT 143
Qy 121 TCAGATCATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Db 144 TCAGATCATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 203
Qy 181 GAGCCATCAAGCTGTCTGCACGCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 240
Db 204 GAGCCATCAAGCTGTCTGCACGCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 263
Qy 241 AGAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCGGAGAGACAGTGTCAAGGCA 300
Db 264 AGAAGATGCCCTGTCTCAGAAAGATGCTGTGGCCCTCGGAGAGACAGTGTCAAGGCA 323
Qy 301 CGGAATCCAGAGCGCAGGCTACGCGCCCGCTCGGCCCGCTCGGCCCGCTGGCCGTGC 360
Db 324 CGGAATCCAGAGCGCAGGCTACGCGCCCGCTCGGCCCGCTCGGCCCGCTGGCCGTGC 383
Qy 361 GCCCTTGGCCCGGAGCGCTTCCACCGCTTCCAGCCCGCTTCCAGCCCGCTTCCAGCGA 420
Db 384 GCCCTTGGCCCGGAGCGCTTCCACCGCTTCCAGCCCGCTTCCAGCCCGCTTCCAGCGA 443
Qy 421 CGAGATGACCTGGCCCGCCACATCTCGCTGTACAGCGGGAGGAGCCCCACCTTACCA 480
Db 444 CGAGATGACCTGGCCCGCCACATCTCGCTGTACAGCGGGAGGAGCCCCACCTTACCA 503
Qy 481 GGGCCCTGACCTTCCAGCTTCCGGACCCCGAGCAGCAGCTTGAACCTGAACCGGAGTC 540
Db 504 GGGCCCTGACCTTCCAGCTTCCGGACCCCGAGCAGCAGCTTGAACCTGAACCGGAGTC 563
Qy 541 GGTGCGGCACCCCGCCAAACAGAACCATCTTCGACAGTACCTGTAGTGGATAGTCCAGGCT 600
Db 564 GGTGCGGCACCCCGCCAAACAGAACCATCTTCGACAGTACCTGTAGTGGATAGTCCAGGCT 623
Qy 601 GGGGGGCGCCCTGCGCCCGCCAGAGTAATCTGGGCATCAGCGCCAGCTGTCTAGCGAGCGG 660
Db 624 GGGGGGCGCCCTGCGCCCGCCAGAGTAATCTGGGCATCAGCGCCAGCTGTCTAGCGAGCGG 683
Qy 661 CGGGCGCATGGA--GGGCGCGCCCGCCACCTACAGCAGGCTCATCGGCCACTACCGGGG 718
Db 684 CGGGCGCATGGAAGGGGGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGG 743
```

```
FEATURES
source
High quality sequence stop: 627.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

```

QY 719 TCCTCTCTCCAGCACCAGC 737
    ||| ||| ||| ||| |||
Db 744 GGGTCCTCTCTTCCAGC 762

RESULT 14
LOCUS B0691705
DEFINITION AGENCOURT_8046876 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6208939
5', mRNA sequence.
ACCESSION B0691705
VERSION B0691705.1 GI:21817021
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 655)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2367 row: 9 column: 20
High quality sequence stop: 645.
Location/Qualifiers
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208939"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 155 a 216 c 193 g 91 t
ORIGIN
Query Match 61.5%; Score 652; DB 13; Length 655;
Best Local Similarity 100.0%; Pred. No. 2.2e-112;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 TACCTGCAGCAGATGACCTCGCGCCACCATCTCGCTGTACAGCGGAGAGGCC 469
Db 1 TACCTGCAGCAGATGACCTCGCGCCACCATCTCGCTGTACAGCGGAGAGGCC 60

QY 470 CCACCTTACCGGGCCCTCGACCTTCAGCTTCGGGACCCCGACGACGCTGGAATG 529
Db 61 CCACCTTACCGGGCCCTTCGACCTTCGAGCTTCGGGACCCCGACGACGCTGGAATG 120

QY 530 AACCGGAGTCTGGTGGCGCACCCGCCAAACAGAACCATCTTCGACAGTCACTGTGGAT 589
Db 121 AACCGGAGTCTGGTGGCGCACCCGCCAAACAGAACCATCTTCGACAGTCACTGTGGAT 180

QY 590 AGTGCAGCTGGGCGGCCCTTCGCCCCACAGTAACCTGGGATACAGCGCCACGTGC 649
Db 181 AGTGCAGCTGGGCGGCCCTTCGCCCCACAGTAACCTGGGATACAGCGCCACGTGC 240

QY 650 TACGCGAGCGGGCGCATGAGGGGCGCGCCGCCACCTACAGCGAGGTCTCGGCCAC 709

```

```

Db 241 TACGCGAGCGGGCGCATGAGGGCGCGCCACCTACAGAGGTTCATCGGCCAC 300
QY 710 TACCGGGGTCTCTTCCAGCAGCAGCAGTGGCCCTCTCTTCTGCTGGAGGG 769
Db 301 TACCGGGGTCTCTTCCAGCAGCAGCAGTGGCCCTCTCTTCTGCTGGAGGG 360
QY 770 ACCGGCTCCACACACATCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 829
Db 361 ACCGGCTCCACACACATCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 420
QY 830 AAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGG 889
Db 421 AAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGG 480
QY 890 TAGGTGAAAAGGAGAACACTCCGGCTTCTTAGAAGAGAGTGAAGAGAGCGGGGG 949
Db 481 TAGGTGAAAAGGAGAACACTCCGGCTTCTTAGAAGAGAGTGAAGAGAGCGGGGG 540
QY 950 CGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGT 1009
Db 541 CGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGT 600
QY 1010 GATGCTGCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 1061
Db 601 GATGCTGCTGAATGCACAGCTTAAGAGAGCTTGCAAAAAA 652

RESULT 15
LOCUS AI761441/c
DEFINITION W65F07.x1 Soares NSF F8_9W OT PA P S1 Homo sapiens cDNA clone
IMAGE:2369989 3. Similar to TR:O15168 O15168 CLONE 22. [3]
TR:O15167 TR:O15165 ; mRNA sequence.
ACCESSION AI761441
VERSION AI761441.1 GI:5177108
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 693)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1105 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 464.
Location/Qualifiers
1. 693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2369989"
/lab_host="DH10B"
/clone_lib="Soares NSF F8_9W OT PA P S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1.
FEATURES
source

```

758280-760583, 772104-774407 Soares N5HPA pool 1:  
304776-306311, 320136-322823, 326280-326663 Soares N5HOT  
pool 1: 723720-726407, 739080-740999 Subtraction by Bento  
Soares and M. Fatima Bonaldo."  
BASE COUNT 101 a 205 c 237 g 149 t 1 others  
ORIGIN

Query Match 60.7%; Score 644.2; DB 9; Length 693;  
Best Local Similarity 98.1%; Pred. No. 6.2e-111;  
Matches 683; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

Qy	359	CGGCCCTTCGGCCACGCGGAGCGCTTCACAGCGCTTCACAGCCACCTATCCGTACCTGCAG	418
Db	693	CGGCCCTTCGGCCACGCGGAGCGCTTCACAGCGCTTCACAGCCACCTATCCGTACCTGCAG	634
Qy	419	CACGAGATCGACCTCGCCGCCACCACTCTCGTGTTCAGACGGGAGGAGCCGCCACCTTAC	478
Db	633	CACGAGATCGA-CTGCCGCCACCACTCTCGTGTTCAGACGGG--AGGAGCCGCCACCTTAC	577
Qy	479	CAGGCCCTTCGACCTTCAGCTTCGGGACCCCGAGCAGCTGGAJCTGAAACCGGAG	538
Db	576	CAGGCCCTTCGACCTTCAGCTTCGGGACCCCGAGCAGCTGGAJCTGAAACCGGAG	517
Qy	539	TCGGTGGCGGCACCCCAACAGACCATCTTCGACAGTGACCTGATGGATAGTCCGAGG	598
Db	516	TCGGTGGCGGCACCCCAACAGACCATCTTCGACAGTGACCTGATGGATAGTCCGAGG	457
Qy	599	CTGGCGGCCCTTCGCCGCCACGAGTAATCTGGGCATCAGCGCCACCTGCTACGGCAGC	658
Db	456	CTGGCGGCCCTTCGCCGCCACGAGTAATCTGGGCATCAGCGCCACCTGCTACGGCAGC	397
Qy	659	GCGGGCGCATGGAGGGCGCGCCCACTACAGCGAGTTCATCGGCCACTTACCCGGGG	718
Db	396	GCGGGCGCATGGAGGGCGCGCCCACTACAGCGAGTTCATCGGCCACTTACCCGGGG	337
Qy	719	TCCTCTTCAGCAGCAGCAGCAGTGGCGCCCTCTTCTGGA-GGGGACCCGGCT	777
Db	336	TCCTCTTCAGCAGCAGCAGCAGTGGCGCCCTCTTCTGGA-GGGGACCCGGGT	277
Qy	778	CCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAA	837
Db	276	CCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAA	217
Qy	838	ACAGAAAGGACACCTCTCTAGGGTCCCAAGGGGGCGGGCTGGGTAGGTGAA	897
Db	216	ACAGAAAGGACACCTCTCTAGGGTCCCAAGGGGGCGGGCTGGGTAGGTGAA	157
Qy	898	AAGCAGAACACTCGCGCTTCTTAGAAGAGAGTGGAGGAGCGGGGGCGGAGCAA	957
Db	156	AAGCAGAACACTCGCGCTTCTTAGAAGAGAGTGGAGGAGCGGGGGCGGAGCAA	97
Qy	958	CGCATCGTGTGGCCCTCCCTCCCACTCCCTGTGTATAATATTTACATGTGTCTG	1017
Db	96	CGCATCGTGTGGCCCTCCCTCCCACTCCCTGTGTATAATATTTACATGTGTCTG	37
Qy	1018	GTCTGAATGCACAAAGCTAAGAGAGCTTGCACAAAAA	1053
Db	36	GTCTGAATGCACAAAGCTAAGAGAGCTTGCACAAAAA	1

Search completed: December 8, 2003, 03:56:09  
Job time : 2647 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 01:56:10 ; Search time 1262 Seconds  
(without alignments)  
2794.246 Million cell updates/sec

Title: US-09-857-826B-44  
Perfect score: 1061  
Sequence: 1 tctccttggttggtgta.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1034.4	97.5	1141	13	US-10-301-822-208
2	1034.4	97.5	1141	15	Sequence 208, App
3	1034.4	97.5	1850	13	Sequence 412, App
4	1034.4	97.5	4527	13	Sequence 44, App
5	1022.4	96.4	1140	13	Sequence 2, Appli
6	1021.2	96.2	1066	14	Sequence 1, Appli
7	961.8	90.7	969	11	Sequence 71, Appl
8	956	90.1	4839	13	Sequence 55, Appl
9	956	90.1	4839	13	Sequence 119, App
10	956	90.1	4839	13	Sequence 84, Appl
11	800.6	75.5	1321	13	Sequence 85, Appl
12	795.2	74.9	806	13	Sequence 1, Appli
13	757.4	71.4	759	13	Sequence 45, Appl
14	752.6	70.9	861	13	Sequence 2, Appli
15	728.8	68.7	1583	15	Sequence 3, Appli
					US-10-000-256A-32

C 16	644.2	60.7	693	10	US-09-934-249-14	Sequence 14, Appl
17	523.8	49.4	878	10	US-09-934-249-12	Sequence 12, Appl
18	521.4	49.1	1713	11	US-09-796-753-57	Sequence 57, Appl
19	350.8	33.1	2170	13	US-10-094-749-197	Sequence 197, App
20	350	33.0	8093	10	US-09-934-249-16	Sequence 16, Appl
21	349	32.9	921	13	US-10-251-598-7	Sequence 0, Appli
22	349	32.9	8065	13	US-10-251-598-6	Sequence 0, Appli
23	332.8	31.4	467	11	US-09-918-995-2074	Sequence 2074, Ap
24	281	26.5	867	13	US-10-251-598-8	Sequence 0, Appli
25	249.2	23.5	475	10	US-09-934-249-15	Sequence 15, Appl
26	218.2	20.6	368	10	US-09-783-590-3464	Sequence 3464, Ap
27	60	5.7	60	13	US-09-908-975-13620	Sequence 13620, A
28	59.2	5.6	65	10	US-09-783-590-3488	Sequence 3488, Ap
C 29	54.8	5.2	5452	13	US-10-017-161-1481	Sequence 1481, Ap
C 30	54.6	5.1	3133	13	US-10-017-161-1483	Sequence 1483, Ap
C 31	53	5.0	1117	13	US-10-017-161-1403	Sequence 1403, Ap
C 32	50.2	4.7	2316	15	US-10-156-761-4416	Sequence 4416, Ap
C 33	50.2	4.7	9025608	15	US-10-156-761-1	Sequence 1, Appli
C 34	49	4.6	15738	13	US-10-329-079-46	Sequence 46, Appl
C 35	49	4.6	61944	13	US-10-329-079-34	Sequence 34, Appl
C 36	48.4	4.6	11304	11	US-09-764-891-5801	Sequence 5801, Ap
C 37	48	4.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
C 38	47.4	4.5	1215	13	US-09-953-348-42	Sequence 42, Appl
C 39	47.4	4.5	1215	15	US-10-267-255-42	Sequence 42, Appl
C 40	47.4	4.5	53500	13	US-09-953-348-76	Sequence 76, Appl
C 41	47.4	4.5	53500	15	US-10-267-255-76	Sequence 76, Appl
C 42	47.2	4.4	1233	15	US-10-166-087-9	Sequence 9, Appli
C 43	47.2	4.4	32539	15	US-10-166-087-1	Sequence 1, Appli
C 44	46.8	4.4	936	15	US-10-156-761-2133	Sequence 2133, Ap
C 45	46.8	4.4	2715	15	US-10-156-761-2245	Sequence 2245, Ap

## ALIGNMENTS

## RESULT 1

US-10-301-822-208  
; Sequence 208, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF COLON CANCER  
; FILE REFERENCE: MPW01-029P2RNM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 208  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: [96]... (854)  
US-10-301-822-208  
Query Match 97.5%; Score 1034.4; DB 13; Length 1141;  
Best Local Similarity 99.2%; Pred. No. 8.6e-274;

Matches 1050; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```
QY 4 TCCTTGGGTTGGGTTGAAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCAGCTGCTGG 63
Db 1 TCCTTGGGTTGGGTTGAAAGCGCTTGGGGTTTC--GTGGCCATGATCCCCAGCTGCTGG 58
QY 64 AGAAGCTGAAGGGGCGAGCTCTCTCGGAAACCCAGGCAATGGCGGAGCTGGAGTTGTTC 123
Db 59 AGAAGCTGAAGGGGCGAGCTCTCTCGGAAACCCAGGCAATGGCGGAGCTGGAGTTGTTC 118
QY 124 GATCATCATCATCTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 183
Db 119 GATCATCATCATCTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 178
QY 184 CCATACAAAGCTGTCTGACGGTCTTTCATGAGCGGCACAGCCAGGCGGAGAGAGA 243
Db 179 CCATACAAAGCTGTCTGACGGTCTTTCATGAGCGGCACAGCCAGGCGGAGAGAGA 238
QY 244 AGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGACAGAGTGTGAGGCAACGG 303
Db 239 AGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGACAGAGTGTGAGGCAACGG 298
QY 304 AATCCAGAGCGCGAGCTTACGCGCCGCTCGGCGCCACCGACCGCTTGGCGTGCGCC 363
Db 299 AATCCAGAGCGCGAGCTTACGCGCCGCTCGGCGCCACCGACCGCTTGGCGTGCGCC 358
QY 364 CTTGCCCGAGCGGAGCGCTTCCAGCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 423
Db 359 CTTGCCCGAGCGGAGCGCTTCCAGCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 418
QY 424 GATGACCTGCGCGCCACCATCTCGTGTGACAGCGGAGGAGCGCCACCTACAGGG 483
Db 419 GATGACCTGCGCGCCACCATCTCGTGTGACAGCGGAGGAGCGCCACCTACAGGG 478
QY 484 CCCCTGACCTTCCAGCTTCCGGGACCCCGAGCAGCAGCTGGAATCAACCGGAGTCGGT 543
Db 479 CCCCTGACCTTCCAGCTTCCGGGACCCCGAGCAGCAGCTGGAATCAACCGGAGTCGGT 538
QY 544 GCGCGACCCCAACAGAACCATCTTTCAGAGTACCTGATGATGATGATGATGATGATG 603
Db 539 GCGCGACCCCAACAGAACCATCTTTCAGAGTACCTGATGATGATGATGATGATGATG 598
QY 604 CGGCGCTTCCGCGCCCGAGCAGTAACTCGGGCATACAGCGCAGCTGCTACGGCAGCGGG 663
Db 599 CGGCGCTTCCGCGCCCGAGCAGTAACTCGGGCATACAGCGCAGCTGCTACGGCAGCGGG 658
QY 664 GCGCATGAGAGGGCGCGCCCGCCACCTACAGCAGAGTCTACGCGCCTACCTACCGGGTCTCT 723
Db 659 GCGCATGAGAGGGCGCGCGCCCGCCACCTACAGCAGAGTCTACGCGCCTACCTACCGGGTCTCT 718
QY 724 CTTCCAGCACCAGCAGAGCAGTGGCGCGCCCTCTTCTGAGAGGGGACCGCGCTCACCA 783
Db 719 CTTCCAGCACCAGCAGAGCAGTGGCGCGCCCTCTTCTGAGAGGGGACCGCGCTCACCA 778
QY 784 CACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGGCAAGAGAGATTAACAGAA 843
Db 779 CACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGGCAAGAGAGATTAACAGAA 838
QY 844 AGGACACCTCTCTAGAGTCCCGAGGGGCGCGGCTGGGGCTGGGTAGGTGAAAGGCA 903
Db 839 AGGACACCTCTCTAGAGTCCCGAGGGGCGCGGCTGGGGCTGGGTAGGTGAAAGGCA 898
QY 904 GAACACTCGCGCTTCTTAGAGAGAGGTGAGAGGAGCGCGGGGCGGAGCAACGCATC 963
Db 899 GAACACTCGCGCTTCTTAGAGAGAGGTGAGAGGAGCGCGGGGCGGAGCAACGCATC 958
QY 964 GTGTGCGCTTCCCTCCACCTCCCTGTGTATTAATTTACATGTGATGCTGTCTGA 1023
Db 959 GTGTGCGCTTCCCTCCACCTCCCTGTGTATTAATTTACATGTGATGCTGTCTGA 1018
QY 1024 ATGCACAAGCTAAGAGAGCTTGCACCAAAAAA 1061
Db 1019 ATGCACAAGCTAAGAGAGCTTGCACCAAAAAA 1056
```

## RESULT 2

US-10-205-823-412  
; Sequence 412, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Wonsley, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 412  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-205-823-412

Query Match 97.5%; Score 1034.4; DB 15; Length 1141;  
Best Local Similarity 99.2%; Pred. No. 8.6e-274;  
Matches 1050; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

```
QY 4 TCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCAGCTGCTGG 63
Db 1 TCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTC--GTGGCCATGATCCCCAGCTGCTGG 58
QY 64 AGAAGCTGAAGGGGCGAGCTCTCTCGGAAACCCAGGCAATGGCGGAGCTGGAGTTGTTC 123
Db 59 AGAAGCTGAAGGGGCGAGCTCTCTCGGAAACCCAGGCAATGGCGGAGCTGGAGTTGTTC 118
QY 124 GATCATCATCATCTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 183
Db 119 GATCATCATCATCTGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 178
QY 184 CCATACAAAGCTGTCTGACGGTCTTTCATGAGCGGCACAGCCAGGCGGAGAGAGA 243
Db 179 CCATACAAAGCTGTCTGACGGTCTTTCATGAGCGGCACAGCCAGGCGGAGAGAGA 238
QY 244 AGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGACAGAGTGTGAGGCAACGG 303
Db 239 AGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCGGAGACAGAGTGTGAGGCAACGG 298
QY 304 AATCCAGAGCGCGAGCTTACGCGCCGCTCGGCGCCACCGACCGCTTGGCGTGCGCC 363
Db 299 AATCCAGAGCGCGAGCTTACGCGCCGCTCGGCGCCACCGACCGCTTGGCGTGCGCC 358
QY 364 CTTGCCCGAGCGGAGCGCTTCCAGCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 423
Db 359 CTTGCCCGAGCGGAGCGCTTCCAGCGCTTCCAGCGCCACCTATCGTACCTGACAGCA 418
```









Db 482 TTACCTGATGATGTCAGGCTGGGCGGCCCTTGCCTCCCGCAGAGTAATCGGGCAT 541  
Qy 637 CAGCGCCACGCTGCTACGGCAGCGCGCGCGCGATGAGGGCGCGCCGCCACCTACAGCGA 696  
Db 542 CAGCGCCACGCTGCTACGGCAGCGCGCGCGCGATGAGGGCGCGCCGCCACCTACAGCGA 601  
Qy 697 GGTCTATCGGCGCACTACCGGGGCTCTCTTCCAGCAGCAGCAGCAGCAGTGGCGCGCCCTC 756  
Db 602 GGTCTATCGGCGCACTACCGGGGCTCTCTTCCAGCAGCAGCAGCAGCAGTGGCGCGCCCTC 661  
Qy 757 CTTGCTGAGGGGACCGGCTCCACACACACACATCGCGCCCTAGAGGCGCAGCCAT 816  
Db 662 CTTGCTGAGGGGACCGGCTCCACACACACACATCGCGCCCTAGAGGCGCAGCCAT 721  
Qy 817 CTGAGCAAGAGAGAGGATAAAGAGGAGCAGCCTCTCTAGGGTCCCGCGGGCGG 876  
Db 722 CTGAGCAAGAGAGAGGATAAAGAGGAGCAGCCTCTCTAGGGTCCCGCGGGCGG 781  
Qy 877 GGCTGGGCTGGGTAGGTGAAAGGAGCAGCCTCTCTAGGGTCCCGCGGGCGG 936  
Db 782 GGCTGGGCTGGGTAGGTGAAAGGAGCAGCCTCTCTAGGGTCCCGCGGGCGG 841  
Qy 937 GGAAGCGGGGCGCGCAGCAGCAGTGTGGGCGCTCCCTCCACCTCCCTGTGTATA 996  
Db 842 GGAAGCGGGGCGCGCAGCAGTGTGGGCGCTCCCTCCACCTCCCTGTGTATA 901  
Qy 997 AATATTTACATGT 1056  
Db 902 AATATTTACATGT 961  
Qy 1057 AAAAA 1061  
Db 962 AAAAA 966

## RESULT 8

US-10-241-220-119  
; Sequence 119, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantzen, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P501081-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 119  
; LENGTH: 4839  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-241-220-119

Query Match 90.1%; Score 956; DB 13; Length 4839;

Best Local Similarity 98.5%; Pred. No. 4e-252;

Matches 965; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 82 TCTCTCGAACCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT 141  
Db 407 TTTGTTCAGAGCATGGAGTACCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT 466  
Qy 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201  
Db 467 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 526

Qy 202 ACGTCTCTTATAGCGCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCTAGA 261  
Db 527 ACGTCTCTTATAGCGCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCTAGA 586  
Qy 262 AGGATGCTGTGGCCCTCGGAGAGCAGATGCTCAGGCAACGGAATCCAGAGCGCAGGT 321  
Db 587 AGGATGCTGTGGCCCTCGGAGAGCAGATGCTCAGGCAACGGAATCCAGAGCGCAGGT 646  
Qy 322 CTAGCGCCCGCTCGGCGCCACCGACCGCTGCGCGCTTGCGCCCTTGCGCCCGGAGCG 381  
Db 647 CTAGCGCCCGCTCGGCGCCACCGACCGCTGCGCGCTTGCGCCCGGAGCGG 706  
Qy 382 CTTTCAACCGCTTCAGCGCCACCTATCTGCTAGTCTGAGAGCAGATGCTGCGCGCCAC 441  
Db 707 CTTTCAACCGCTTCAGCGCCACCTATCTGCTAGTCTGAGAGCAGATGCTGCGCGCCAC 766  
Qy 442 CATCTCGCTGTACAGCGGGAGGAGCCCGACCTTACCAGGGCCCTGACCCCTCCAGCT 501  
Db 767 CATCTCGCTGTACAGCGGGAGGAGCCCGACCTTACCAGGGCCCTGACCCCTCCAGCT 826  
Qy 502 TCGGAGACCGCGAGCAGCAGCTGGAACCTGAAACCGGGAGTCCGTGCGCGCACCCCAAGAG 561  
Db 827 TCGGAGACCGCGAGCAGCAGCTGGAACCTGAAACCGGGAGTCCGTGCGCGCACCCCAAGAG 886  
Qy 562 AACCATCTTTCAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621  
Db 887 AACCATCTTTCAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 946  
Qy 622 CAGTAATCTCGGGCATAGCGCCAGTGTCTAGCGCAGCGGGCGGCGCATGAGAGGGCGGCC 681  
Db 947 CAGTAATCTCGGGCATAGCGCCAGTGTCTAGCGCAGCGGGCGGCGCATGAGAGGGCGGCC 1006  
Qy 682 GCCCACCTTACAGCAGGCTCATCGGCCACTACCCCGGGTCTCTTCCAGCAGCAGCAGAG 741  
Db 1007 GCCCACCTTACAGCAGGCTCATCGGCCACTACCCCGGGTCTCTTCCAGCAGCAGCAGAG 1065  
Qy 742 CAGTGGCGCGCCCTCTTGTGGAGGGACCCCGCTCCACACACACATCGCGGCCCT 801  
Db 1067 CAGTGGCGCGCCCTCTTGTGGAGGGACCCCGCTCCACACACACATCGCGGCCCT 1126  
Qy 802 AGAGAGCGCGCATCTGAGCAAGAGAGAGATTAACAGAGACACCTCTCTTAGGG 861  
Db 1127 AGAGAGCGCGCATCTGAGCAAGAGAGAGATTAACAGAGAGACACCTCTCTTAGGG 1186  
Qy 862 TCCCCAGGGGGCGCGCTGGCGCTGCGTAGTGAAGAGGAGAGAGAGAGAGAGAGAGAGAG 921  
Db 1187 TCCCCAGGGGGCGCGCTGGCGCTGCGTAGTGAAGAGGAGAGAGAGAGAGAGAGAGAG 1246  
Qy 922 AGAAGAGAGTGAAG 981  
Db 1247 AGAAGAGAGTGAAG 1306  
Qy 982 ACCTCCCTGTGTATAAATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1041  
Db 1307 ACCTCCCTGTGTATAAATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1366  
Qy 1042 CTTGCAAAAAA 1061  
Db 1367 CTTGCAAAAAA 1386

## RESULT 9

US-10-269-909-84

; Sequence 84, Application US/10269909

; Publication No. US20030180747A1

; GENERAL INFORMATION:

; APPLICANT: HRUBAN, RALPH H.

; APPLICANT: ARGANI, PEDRAM

; APPLICANT: IACOBUIO-DONAHUE, CHRISTINE

; APPLICANT: MALTRA, ANIRBAN

; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES

; FILE REFERENCE: 58303(71699)

; CURRENT APPLICATION NUMBER: US/10/269,909



; CURRENT FILING DATE: 2003-10-11  
; PRIOR APPLICATION NUMBER: 60/328,609  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 60/332,754  
; PRIOR FILING DATE: 2001-11-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 84  
; LENGTH: 4839  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-269-909-84

Query Match 90.1%; Score 956; DB 13; Length 4839;  
Best Local Similarity 98.5%; Pred. No. 4e-252;  
Matches 965; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 82 TCTCTGCGAAACACGAGGCTGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141  
DB 407 TTTGTTCCAGAGCATGGAGTACCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 466  
QY 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201  
DB 467 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 526  
QY 202 ACGGTCTTCATCAGCCGACAGCCAGGCGGAGGAGAGAGATGCCCTGTCTCAGA 261  
DB 527 ACGGTCTTCATCAGCCGACAGCCAGGCGGAGGAGAGAGATGCCCTGTCTCAGA 586  
QY 262 AGGATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 321  
DB 587 AGGATGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 646  
QY 322 CTACGCGCGCTTCAGCGCCACCGTGTACGCGAGCGGCGGAGGAGAGATGCCCTGTCTCAGA 441  
DB 707 CTTCACCGCTTCAGCGCCACCGTGTACGCGAGCGGAGGAGAGATGCCCTGTCTCAGA 766  
QY 442 CATCTCGCTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 501  
DB 767 CATCTCGCTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 826  
QY 502 TCGGAGCCCGAGCAGAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGGAGGAG 561  
DB 827 TCGGAGCCCGAGCAGAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGGAGGAG 886  
QY 562 AACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGGCGGCGGCGGCGGCGG 621  
DB 887 AACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGGCGGCGGCGGCGGCGG 946  
QY 622 CAGTAACTCGGCGCATCAGCGCCACCGTGTACGCGAGCGGCGGCGGCGGCGGCGG 681  
DB 947 CAGTAACTCGGCGCATCAGCGCCACCGTGTACGCGAGCGGCGGCGGCGGCGGCGG 1006  
QY 682 GCCCACTTACAGCGAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCAGCAG 741  
DB 1007 GCCCACTTACAGCGAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCAGCAG 1066  
QY 742 CAGTGGGCGGCGCTCTTGTGGAGGGGACCGGCTCCACCAACACACATCGGCGGCGCT 801  
DB 1067 CAGTGGGCGGCGCTCTTGTGGAGGGGACCGGCTCCACCAACACACATCGGCGGCGCT 1126  
QY 802 AGAGAGCGCAGCATCTGGAGCAAAAGAGAGATGAAACAGAAAGGACACCTCTCTAGGG 861  
DB 1127 AGAGAGCGCAGCATCTGGAGCAAAAGAGAGATGAAACAGAAAGGACACCTCTCTAGGG 1186  
QY 862 TCCCAAGGGGGCGGGCTGGGGCTGGGTAGTGAAGAGGAGAGACACTCTCGCGCTTCTT 921  
DB 1187 TCCCAAGGGGGCGGGCTGGGGCTGGGTAGTGAAGAGGAGAGACACTCTCGCGCTTCTT 1246

QY 922 AGAAGAGAGTGAAGAGGCGGGGCGGAGCAAGCATCTGTGTGGCCCTCCCTCCC 981  
DB 1247 AGAAGAGAGTGAAGAGGCGGGGCGGAGCAAGCATCTGTGTGGCCCTCCCTCCC 1306  
QY 982 ACCTCCCTGTGTATAAATATTTTACATGTGATCTGTGTCTGAATGCAACAGCTAAGAGAG 1041  
DB 1307 ACCTCCCTGTGTATAAATATTTTACATGTGATCTGTGTCTGAATGCAACAGCTAAGAGAG 1366  
QY 1042 CTTGCAAAAAA 1061  
DB 1367 CTTGCAAAAAA 1386

RESULT 10  
US-10-269-909-85  
; Sequence 85, Application US/10269909  
; Publication No. US20030180747A1  
; GENERAL INFORMATION:  
; APPLICANT: HRUBAN, RALPH H.  
; APPLICANT: ARGANI, PEDRAM  
; APPLICANT: IACOBUIZIO-DONAHUE, CHRISTINE  
; APPLICANT: MAITRA, ANIRBAN  
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES  
; FILE REFERENCE: 58303(71699)  
; CURRENT APPLICATION NUMBER: US/10/269,909  
; CURRENT FILING DATE: 2003-10-11  
; PRIOR APPLICATION NUMBER: 60/328,609  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 60/332,754  
; PRIOR FILING DATE: 2001-11-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 85  
; LENGTH: 4839  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-269-909-85

Query Match 90.1%; Score 956; DB 13; Length 4839;  
Best Local Similarity 98.5%; Pred. No. 4e-252;  
Matches 965; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 82 TCTCTCGAAACACGAGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141  
DB 407 TTTGTTCCAGAGCATGGAGTACCGAGCTGGAGTTTGTTCAGATCATCATCGTGT 466  
QY 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201  
DB 467 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 526  
QY 202 ACGGTCTTCATCAGCGCGCACAGCCAGGCGGAGGAGAGAGATGCCCTGTCTCAGA 261  
DB 527 ACGGTCTTCATCAGCGCGCACAGCCAGGCGGAGGAGAGAGATGCCCTGTCTCAGA 586  
QY 262 AGGATGCTGTGGCGCTTCGAGAGCAGAGTGTCAAGGCAACGGAATCCAGAGCGCAGGT 321  
DB 587 AGGATGCTGTGGCGCTTCGAGAGCAGAGTGTCAAGGCAACGGAATCCAGAGCGCAGGT 646  
QY 322 CTACGCGCGCTTCGCGCCACCGAGCCGCTGGCGGCTGGCGGCTTCGCGGCGGAGCG 381  
DB 647 CTACGCGCGCTTCGCGCCACCGAGCCGCTGGCGGCTTCGCGGCGGAGCG 706  
QY 382 CTTTCAACGCTTCAGCGCCACCTATCGTACCTGAGCAGCAGATCGACCTGCGGCCAC 441  
DB 707 CTTTCAACGCTTCAGCGCCACCTATCGTACCTGAGCAGCAGATCGACCTGCGGCCAC 766  
QY 442 CATCTCGCTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 501  
DB 767 CATCTCGCTGTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 826  
QY 502 TCGGAGCCCGAGCAGAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGGAGGAG 561  
DB 827 TCGGAGCCCGAGCAGAGCTGGAATGAAACCGGAGTGGTGGCGGACCCCGGAGGAG 886





```
; SEQ ID NO 45
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-241-220-45

Query Match      74.9%; Score 795.2; DB 13; Length 806;
Best Local Similarity 99.6%; Pred. No. 3.5e-208;
Matches 797; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 75 CGGACGGTCTCTCGGAACACGAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCA 134
Db 7 CGGACAGTCTCTCGGAACACGAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCA 66

QY 135 TCGTGGTGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 194
Db 67 TCGTGGTGTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126

QY 195 TGTCTGCACGGTCTCTCATCGCCGCGCACAGCCAGCGGGCGGAGGAGAGATGCCCTGT 254
Db 127 TGTCTGCACGGTCTCTCATCGCCGCGCACAGCCAGCGGGCGGAGGAGAGATGCCCTGT 186

QY 255 CCTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCAGAGC 314
Db 187 CCTCAGAGGATGCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCAGAGC 246

QY 315 CGCAGGTCTACGCCCGGCTCGGCCACCGACCGGCTGGCCGTGGCGGCTTGGCCGAGC 374
Db 247 CGCAGGTCTACGCCCGGCTCGGCCACCGACCGGCTGGCCGTGGCGGCTTGGCCGAGC 306

QY 375 GGGAGCGTCTCACCGCTTCCAGCCACCTATCTCGTACCTGACGACGAGATCGACCTGC 434
Db 307 GGGAGCGTCTCACCGCTTCCAGCCACCTATCTCGTACCTGACGACGAGATCGACCTGC 366

QY 435 CGCCACCATCTCGTGTGACAGCGGGAGGAGCCGCCACCTTACAGGGGCCCTCGACCC 494
Db 367 CGCCACCATCTCGTGTGACAGCGGGAGGAGCCGCCACCTTACAGGGGCCCTCGACCC 426

QY 495 TCCAGTTCGGACCCCGGAGGAGAGCTGGAACTGAACTGGAGTGGTGGCGGACCC 554
Db 427 TCCAGTTCGGACCCCGGAGGAGAGCTGGAACTGAACTGGAGTGGTGGCGGACCC 486

QY 555 CAAACAGAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGAT 614
Db 487 CAAACAGAACCATCTTCGACAGTACCTGATGATGATGATGATGATGATGATGATGAT 546

QY 615 CCCCCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 674
Db 547 CCCCCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 606

QY 675 GGGCGCGGCCACCTACAGCGAGGTGATCGGCCACTACCGGGGTCTCTTCCAGGACC 734
Db 607 GGGCGCGGCCACCTACAGCGAGGTGATCGGCCACTACCGGGGTCTCTTCCAGGACC 666

QY 735 AGCAGAGCAGTGGCGGCGCTCTTCTGAGGGGACCGCGCTCCACACACACACATCG 794
Db 667 AGCAGAGCAGTGGCGGCGCTCTTCTGAGGGGACCGCGCTCCACACACACACATCG 726

QY 795 CGCCCTCTAGAGCGCAGCCATCTGGAGCAAGAGAGATGATGATGATGATGATGATGAT 854
Db 727 CGCCCTCTAGAGCGCAGCCATCTGGAGCAAGAGAGATGATGATGATGATGATGATGAT 786

QY 855 TCTAGGGTCCCGAGGGGGC 874-
Db 787 TCTAGGGTCCCGAGAGGGC 806
```

## RESULT 13

```
US-10-390-045-2
; Sequence 2, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
```

```
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: POYNUCLEOTIDE ARRAY
; CURRENT APPLICATION NUMBER: US/10/390,045
; PRIORITY FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 2
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-390-045-2
```

```
Query Match      71.4%; Score 757.4; DB 13; Length 759;
Best Local Similarity 99.9%; Pred. No. 8.1e-198;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 101 ATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 160
Db 1 ATGCGGAGCTGGAGTTTGTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 60

QY 161 GTGTGTATCATCGTCTGAGCCACTTACAGCTGTCTGACGCTCTTCCATCAGCCGG 220
Db 61 GTGTGTATCATCGTCTGAGCCACTTACAGCTGTCTGACGCTCTTCCATCAGCCGG 120

QY 221 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCG 280
Db 121 CACAGCCAGGGCGGAGGAGAGATGCCCTGTCTCAGAGGATGCTGTGGCCCTCG 180

QY 281 GAGAGCAGTGTGAGGCAACGGAATCCAGAGCCGACAGTCTACGCCGCCCTCGGCC 340
Db 181 GAGAGCAGTGTGAGGCAACGGAATCCAGAGCCGACAGTCTACGCCGCCCTCGGCC 240

QY 341 ACCGACCGCTGGCGGCTGCGCCCTTGGCCAGCGGAGCGCTTCCACCGCTTCCAGCC 400
Db 241 ACCGACCGCTGGCGGCTGCGCCCTTGGCCAGCGGAGCGCTTCCACCGCTTCCAGCC 300

QY 401 ACCTATCCGTACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGCTGTGACAGCGG 460
Db 301 ACCTATCCGTACCTGCAGCAGAGATCGACCTGCGGCCACCATCTCGCTGTGACAGCGG 360

QY 461 GAGGAGCCCCACCTTACAGGGCCCTTGACCTCTGACCTTGGGAGCCCCGAGGAGCAG 520
Db 361 GAGGAGCCCCACCTTACAGGGCCCTTGACCTCTGACCTTGGGAGCCCCGAGGAGCAG 420

QY 521 CTGGAATCTGAACCGGAGTGGTGGCGCACCCCAACACAGAACATCTTCGACAGTCA 580
Db 421 CTGGAATCTGAACCGGAGTGGTGGCGCACCCCAACACAGAACATCTTCGACAGTCA 480

QY 581 CTGATGATAGTGCAGGCTGGGCGGCCCTTGCCTGCCCCCAGCAGTAACTCGGGCATCAGC 640
Db 481 CTGATGATAGTGCAGGCTGGGCGGCCCTTGCCTGCCCCCAGCAGTAACTCGGGCATCAGC 540

QY 641 GCCAGTGTACGGCAGCGGGCGGCGATGAGGGGCGCGCCGACCTACAGCGAGGTC 700
Db 541 GCCAGTGTACGGCAGCGGGCGGCGATGAGGGGCGCGCCGACCTACAGCGAGGTC 600

QY 701 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGCAGTGGGCGGCCCTCTTGT 760
Db 601 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGAGCAGTGGGCGGCCCTCTTGT 660

QY 761 CTGGAGGGGAGCCCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGCAGCCTCTGG 820
Db 661 CTGGAGGGGAGCCCGGCTCCACACACACATCGCGGCCCTTAGAGAGCGCAGCCTCTGG 720
```

Query Match	68.7%;	Score 728.8;	DB 15;	Length 1583;
Best Local Similarity	98.9%;	Pred. No. 6.8e-190;		
Matches	744;	Conservative	0;	Mismatches 7; Indels 1; Gaps 1;
QY	311	GAGCGGAGGTTCTACGCCCGCCGCTCGGCCCCACCGACCGCCTGGCCGTGGCCGCTTCGCC	370	
Db	825	GCGCCGAGGTTCTACGCCCGCCGCTCGGCCCCACCGACCGCCTGGCCGTGGCCGCTTCGCC	884	
QY	371	CAGCGGAGGCTTCCACACGCTTTCAGGCCACACTATCCGTACCTGTCAGACACGAGATCGAC	430	
Db	885	CAGCGGAGGCTTTCACACGCTTTCAGGCCACACTATCCGTACCTGTCAGACACGAGATCGAC	944	
QY	431	CTGCCGCCACCATCTCGCTGTTCAGACGGGGAGGACCCACCTACCAAGGGCCCTTCG	490	
Db	945	CTGCCACCCACCATCTCGCTGTTCAGACGGGGAGGACCCCACTTACCAAGGGCCCTTCG	1004	
QY	491	ACCTCTCAGTTTCGGGACCCCGAGCAGCTGGAACTGAAACCGGGAGTCTGGTGGCGCA	550	
Db	1005	ACCTCTCAGTTTCGGGACCCCGAGCAGCTGGAACTGAAACCGGGAGTCTGGTGGCGCA	1064	
QY	551	CCCCAAAACAGAACCATCTTCGACAGTGCCTCATGGATAGTGCCAGGCTGGGGGGCCCC	610	
Db	1065	CCCCAAAACAGAACCATCTTCGACAGTGCCTCATGGATAGTGCCAGGCTGGGGGGCCCC	1124	
QY	611	TGCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCGTGTCTACGGCAGCGCGGCGCATG	670	
Db	1125	TGCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCGTGTCTACGGCAGCGCGGCGCATG	1184	
QY	671	GAGGGCGCGGCCCACTTACAGCGAGGTTCATCGGCCACTACCGGGGTCTCTCTTCAG	730	
Db	1185	GAGGGCGCGGCCCACTTACAGCGAGGTTCATCGGCCACTACCGGGGTCTCTCTTCAG	1244	

Qy 731 CACCAGCAGAGCAGTGGCGCGCCCTCTTGTGGAGGGACCCGGCTCCACACACAC 790  
Db |||||  
Qy 1245 CACCAGCAGAGCAGTGGCGCGCCCTCTTGTGGAGGGACCCGGCTCCACACACAC 1304  
Db |||||  
Qy 791 ATCGCGCCCTTAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACAC 850  
Db |||||  
Qy 1305 ATCGCGCCCTTAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACAC 1364  
Db |||||  
Qy 851 CCTCTTAGGTCCCGAGGGGGCGGGCTGGGCTGGTGGTGAAGAGGAGACACT 910  
Db |||||  
Qy 1365 CCTCTTAGGTCCCGAGGGGGCGGGCTGGGCTGGTGGTGAAGAGGAGACACT 1424  
Db |||||  
Qy 911 CGCGCTTCTTAGAGAGAGTGAGAGGAAGCGGGCGCAGCAACGCAT-CGTGTGG 969  
Db |||||  
Qy 1425 CGCGCTTCTTAGAGAGAGTGAGAGGAAGCGGGCGCAGCAACGCATCGCATGCGTGG 1484  
Db |||||  
Qy 970 CCTCCCCCTCCACCTCCCTGTGTATAATTTACATGTGATGTCTGGTCTGAATGCAC 1029  
Db |||||  
Qy 1485 CCTCCCCCTCCACCTCCCTGTGTATAATTTACATGTGATGTCTGGTCTGAATGCAC 1544  
Db |||||  
Qy 1030 AAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1061  
Db |||||  
Qy 1545 AAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1576  
Db |||||

Search completed: December 8, 2003, 04:18:49  
Job time : 1274 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:50:19 ; Search time 31 Seconds  
(without alignments)  
1511.867 Million cell updates/sec

Title: US-09-857-826B-17  
Perfect score: 1348  
Sequence: 1 MALEFVQIIIVVMVMV.....PLESAIWSKDKQKHPL 252

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US04\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US03\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US02\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US01\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US00\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	252	11 US-09-796-753-56	Sequence 56, Appl
2	1348	100.0	252	12 US-10-241-220-100	Sequence 100, Appl
3	1348	100.0	252	12 US-10-241-220-101	Sequence 101, Appl
4	1348	100.0	252	12 US-10-301-822-209	Sequence 209, Appl
5	1348	100.0	252	12 US-09-821-812-3	Sequence 3, Appl
6	1348	100.0	252	12 US-10-390-045-3	Sequence 3, Appl
7	1348	100.0	252	15 US-10-205-823-413	Sequence 413, Appl
8	1340	99.4	287	10 US-09-934-249-2	Sequence 2, Appl
9	1340	99.4	287	12 US-10-241-220-120	Sequence 120, Appl
10	1332	98.8	249	12 US-10-390-045-11	Sequence 11, Appl
11	1172.5	87.0	274	10 US-09-934-249-13	Sequence 13, Appl
12	1066.5	79.1	217	11 US-09-796-753-58	Sequence 58, Appl
13	1065.5	79.0	241	15 US-10-000-236A-158	Sequence 158, Appl
14	845.5	62.7	244	12 US-10-390-045-12	Sequence 12, Appl
15	845.5	62.7	269	12 US-10-094-749-1836	Sequence 1836, Appl

Sequence 0, Appl  
Sequence 0, Appl  
Sequence 134, Appl  
Sequence 57, Appl  
Sequence 134, Appl  
Sequence 197, Appl  
Sequence 197, Appl  
Sequence 140, Appl  
Sequence 236, Appl  
Sequence 5, Appl  
Sequence 22, Appl  
Sequence 273, Appl  
Sequence 273, Appl  
Sequence 14, Appl  
Sequence 110, Appl  
Sequence 8, Appl  
Sequence 156, Appl  
Sequence 348, Appl  
Sequence 20, Appl  
Sequence 278, Appl  
Sequence 324, Appl  
Sequence 333, Appl  
Sequence 63, Appl  
Sequence 331, Appl  
Sequence 61, Appl  
Sequence 61, Appl  
Sequence 14, Appl  
Sequence 4553, Appl

## ALIGNMENTS

RESULT 1  
US-09-796-753-56  
; Sequence 56, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23

684280-17

;; PRIOR APPLICATION NUMBER: 09/474,071  
;; PRIOR FILING DATE: 1999-12-29  
;; PRIOR APPLICATION NUMBER: 09/474,072  
;; PRIOR FILING DATE: 1999-12-29  
;; PRIOR APPLICATION NUMBER: 09/514,010  
;; PRIOR FILING DATE: 2000-02-25  
;; PRIOR APPLICATION NUMBER: 09/516,745  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 09/572,002  
;; PRIOR FILING DATE: 2000-05-14  
;; PRIOR APPLICATION NUMBER: 09/597,993  
;; PRIOR FILING DATE: 2000-06-19  
;; PRIOR APPLICATION NUMBER: 09/599,596  
;; PRIOR FILING DATE: 2000-06-22  
;; PRIOR APPLICATION NUMBER: 09/630,334  
;; PRIOR FILING DATE: 2000-07-31  
;; PRIOR APPLICATION NUMBER: 09/606,565  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/606,317  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/665,666  
;; PRIOR FILING DATE: 2000-09-20  
;; PRIOR APPLICATION NUMBER: 09/677,751  
;; PRIOR FILING DATE: 2000-09-30  
;; NUMBER OF SEQ ID NOS: 162  
;; SEQ ID NO 56  
;; LENGTH: 252  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-796-753-56

Query Match 100.0%; Score 1348; DB 11; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111; Indels 0; Gaps 0;  
Matches 252; Conservative 0; Mismatches 0;  
QY 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
Db 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
QY 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
QY 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180  
Db 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180  
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 2  
US-10-241-220-100  
; Sequence 100, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220

;; CURRENT FILING DATE: 2002-12-13  
;; NUMBER OF SEQ ID NOS: 120  
;; SEQ ID NO 100  
;; LENGTH: 252  
;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-10-241-220-100  
Query Match 100.0%; Score 1348; DB 12; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111; Indels 0; Gaps 0;  
Matches 252; Conservative 0; Mismatches 0;  
QY 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
Db 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
QY 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
QY 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180  
Db 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180  
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
QY 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 3  
US-10-241-220-101  
; Sequence 101, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 101  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-101

Query Match 100.0%; Score 1348; DB 12; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111; Indels 0; Gaps 0;  
Matches 252; Conservative 0; Mismatches 0;  
QY 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
Db 1 MAELFVQIIIIIVVMVMMVVITCLSHYKLSARSFISRHSGRRRDLSSSEGLWPS 60  
QY 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEPVYAPRPTDRLAVPPFAQRFRHFPQTPYVYLQHEIDLPTTISLSDG 120  
QY 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180  
Db 121 EEPYVQGCCTQLQDRPEQLELNRESVRAPNRTIFDSDMLDMSARLGCPCCPSNSGIS 180

QY 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
Db |||||  
QY 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
Db |||||  
QY 241 SKEKDOKQGHPL 252  
Db 241 SKEKDOKQGHPL 252

RESULT 4  
US-10-301-822-209  
; Sequence 209 Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: BURGART, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF COLON CANCER  
; FILE REFERENCE: MP01-02422RNM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 209  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-301-822-209

Query Match 100.0%; Score 1348; DB 12; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELEFVQIIIVVMMVWVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
Db |||||  
QY 61 ESTVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVYLQHEIDLPTISLSDG 120  
QY 121 EBPYPYQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGTS 180  
Db 121 EBPYPYQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGTS 180  
QY 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
Db 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
QY 241 SKEKDOKQGHPL 252  
Db 241 SKEKDOKQGHPL 252

RESULT 5  
US-09-821-812-3  
; Sequence 3, Application US/09821812  
; Publication No. US20030166520A1  
; GENERAL INFORMATION:

; APPLICANT: Lin, Biaoyang  
; TITLE OF INVENTION: Androgen Regulated Prostate Specific  
; FILE REFERENCE: P-IS 4373  
; CURRENT APPLICATION NUMBER: US/09/821,812  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-821-812-3

Query Match 100.0%; Score 1348; DB 12; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELEFVQIIIVVMMVWVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
Db |||||  
QY 61 ESTVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVYLQHEIDLPTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVYLQHEIDLPTISLSDG 120  
QY 121 EBPYPYQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGTS 180  
Db 121 EBPYPYQGPCTQLQDRPEQQLNRESVRAPPNRTIFDSLDMSARLGGPCPPSSNSGTS 180  
QY 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
Db 181 ATCYSGGMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLLEGTRLHHTHIAPLSAI 240  
QY 241 SKEKDOKQGHPL 252  
Db 241 SKEKDOKQGHPL 252

RESULT 6  
US-10-390-045-3  
; Sequence 3, Application US/10390045  
; Publication No. US20030170713A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: 04995,0057-00000  
; CURRENT APPLICATION NUMBER: US/10/390,045  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US/09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-390-045-3

Query Match 100.0%; Score 1348; DB 12; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9.4e-111;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAELEFVQIIIVVMMVWVITCLSHYKLSARSFISRHSGRRRDLSSGCLWPS 60  
Db |||||

Db 1 MAELFVQIIIIIVVMMVVVITCLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60  
Qy 61 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 120  
Qy 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
Db 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 7  
US-10-205-823-413  
; Sequence 413, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Wensey, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205,823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307,982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314,356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325,020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341,746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362,158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 413  
; LENGTH: 252  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-205-823-413

Query Match 100.0%; Score 1348; DB 15; Length 252;  
Best Local Similarity 100.0%; Pred. No. 9,4e-111;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAELFVQIIIIIVVMMVVVITCLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60  
Db 1 MAELFVQIIIIIVVMMVVVITCLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60  
Qy 61 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 120  
Db 61 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 120  
Qy 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180

Db 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
Qy 181 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 181 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Qy 241 SKEKDQKQGHPL 252  
Db 241 SKEKDQKQGHPL 252

RESULT 8  
US-09-934-249-2  
; Sequence 2, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Richard T.  
; APPLICANT: Landschulz, Katherine T.  
; APPLICANT: Tuti, Thomas G.  
; APPLICANT: Thompson, John F.  
; APPLICANT: Kennedy, Scott P.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
; FILE REFERENCE: P0738/7001/ERP/KA  
; CURRENT APPLICATION NUMBER: US/09/934,249  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/227,159  
; PRIOR FILING DATE: 2000-08-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-09-934-249-2

Query Match 99.4%; Score 1340; DB 10; Length 287;  
Best Local Similarity 99.2%; Pred. No. 5.6e-110;  
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MAELFVQIIIIIVVMMVVVITCLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60  
Db 36 ITELEFVQIIIIIVVMMVVVITCLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 95  
Qy 61 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 120  
Db 96 ESTVSGNGIPEQVYAPPRPTDLAVPPPAQRERFHRQPTTPYLOHEIDLPTTISLSDG 155  
Qy 121 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
Db 156 BEPPPYQGCTQLQRLDPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 215  
Qy 181 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
Db 216 ATCYGSGGRMEGPPPTTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 275  
Qy 241 SKEKDQKQGHPL 252  
Db 276 SKEKDQKQGHPL 287

RESULT 9  
US-10-241-220-120  
; Sequence 120, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas



APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P50101-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 120  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-120

Query Match 99.4%; Score 1340; DB 12; Length 287;  
Best Local Similarity 99.2%; Pred. No. 5.6e-110;  
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MALEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 60  
DB 36 ITELEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 95  
QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRRFHRFQTPYPLQHEIDLPTTISLSDG 120  
DB 96 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRRFHRFQTPYPLQHEIDLPTTISLSDG 155  
QY 121 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
DB 156 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 215  
QY 181 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 240  
DB 216 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 275  
QY 241 SKEKDQKQGHPL 252  
DB 276 SKEKDQKQGHPL 287

RESULT 10  
US-10-390-045-11  
; Sequence 11, Application US/10390045  
; Publication No. US20030170713A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: MOUL, JUDD W.  
; APPLICANT: XU, LINDA L.  
; APPLICANT: SEGAWA, TAKEHIKO  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/10/390,045  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US/09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-390-045-11

Query Match 98.8%; Score 1332; DB 12; Length 249;  
Best Local Similarity 100.0%; Pred. No. 2.4e-109;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 AELEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 61  
DB 1 AELEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 60

QY 62 STVSGNGIPEQVYAPRPTDRLAVPPFAQRRFHRFQTPYPLQHEIDLPTTISLSDG 121  
DB 61 STVSGNGIPEQVYAPRPTDRLAVPPFAQRRFHRFQTPYPLQHEIDLPTTISLSDG 120  
QY 122 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 181  
DB 121 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
QY 182 TCYSGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 241  
DB 181 TCYSGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 240  
QY 242 KEKDQKQGH 250  
DB 241 KEKDQKQGH 249  
RESULT 11  
US-09-934-249-13  
; Sequence 13, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Richard T.  
; APPLICANT: Landschulz, Katherine T.  
; APPLICANT: Turi, Thomas G.  
; APPLICANT: Thompson, John F.  
; APPLICANT: Kennedy, Scott P.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
; FILE REFERENCE: P0738/7001/ERP/KA  
; CURRENT APPLICATION NUMBER: US/09/934,249  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/227,159  
; PRIOR FILING DATE: 2000-08-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Mus Musculus  
US-09-934-249-13

Query Match 87.0%; Score 1172.5; DB 10; Length 274;  
Best Local Similarity 88.1%; Pred. No. 2.9e-95;  
Matches 222; Conservative 11; Mismatches 10; Indels 9; Gaps 3;  
QY 1 MALEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 60  
DB 32 ITELEFVQIIIVVMVMVITCLLSHYKLSARSFISRHSGQRRRDLSSGCLWPS 91  
QY 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRRFHRFQTPYPLQHEIDLPTTISLSDG 120  
DB 92 ESTVSG-GMPEQVYAPRPTDRLAVPPFIORS---RQTPYPLQHEIALPTTISLSDG 147  
QY 121 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 180  
DB 148 EPPPYQGPCTQLQDRPEQQLNRESVRAPNRTIFDSLDMSARLGCGPCPPSSNSGIS 207  
QY 181 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 240  
DB 208 ATCYGSGGRMEGPPPTTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAIWS 262  
QY 241 SKEKDQKQGHPL 252  
DB 263 SKEKDQKQGHPL 274

RESULT 12  
US-09-796-753-58  
; Sequence 58, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:

APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 09/474,071  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/474,072  
; PRIOR FILING DATE: 1999-12-29  
; PRIOR APPLICATION NUMBER: 09/514,010  
; PRIOR FILING DATE: 2000-02-25  
; PRIOR APPLICATION NUMBER: 09/516,745  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 09/572,002  
; PRIOR FILING DATE: 2000-05-14  
; PRIOR APPLICATION NUMBER: 09/597,993  
; PRIOR FILING DATE: 2000-06-19  
; PRIOR APPLICATION NUMBER: 09/599,596  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/630,334  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 09/606,565  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/606,317  
; PRIOR FILING DATE: 2000-06-29  
; PRIOR APPLICATION NUMBER: 09/665,666  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: 09/677,751  
; PRIOR FILING DATE: 2000-09-30  
; NUMBER OF SEQ ID NOS: 162  
; SEQ ID NO 58  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-796-753-58

Query Match 79.1%; Score 1066.5; DB 11; Length 217;  
Best Local Similarity 88.9%; Pred. No. 4.8e-86;  
Matches 201; Conservative 7; Mismatches 9; Indels 9; Gaps 3;  
QY 27 LSHYKLSARSFTRSRQRRRDLSSGCLWPSESTVSGNGIPSEVAPPRPTDLAV 86  
DB 1 LSHYKLSARSFTRSRQRRRDLSSGCLWPSESTVSG-GMPPEVQVYAPPRPTDLAV 59  
QY 87 BPAQRERHRTQPTVYVQLQHEIDLPTTISLSDGEEPPYQGPCTQLQLRDPQQLNRE 146

Db 60 PPFQIORS---RFQPTVYVQLQHEIALPPTISLSDGEEPPYQGPCTQLQLRDPQQLNRE 116  
QY 147 SVRAPPPRTIFDSDLMDSARLGGPCPPSSNSGIGATCYGSGRMEGPPPTTYSEVIGHYGP 206  
Db 117 SVRAPPPRTIFDSDLMDSARLGGPCPPSSNSGIGATCYGSGRMEGPPPTTYSEVIGHYGP 176  
QY 207 SSFOHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKOKGHPL 252  
Db 177 SSFOHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKOKGHPL 217  
RESULT 13  
US-10-000-256A-158  
; Sequence 158, Application US/10000256A  
; Publication No. US20030039983A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Chen, Sei-yu  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0259  
; CURRENT APPLICATION NUMBER: US/10/000,256A  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/244,782  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 240  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 158  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-000-256A-158

Query Match 79.0%; Score 1065.5; DB 15; Length 241;  
Best Local Similarity 82.9%; Pred. No. 6.6e-86;  
Matches 199; Conservative 0; Mismatches 0; Indels 41; Gaps 1;  
QY 54 EGCLWPSESTVSGNGIPE-----P 72  
Db 1 EGCLWPSESTVSGNGIPECCWDPCCRRSAPCAPGSPALCSLHTGARTLPLFGGGRP 60  
QY 73 QVYAPPRTDLAVPPPAQRERHRTQPTVYVQLQHEIDLPTTISLSDGEEPPYQGPCTL 132  
Db 61 QVYAPPRTDLAVPPPAQRERHRTQPTVYVQLQHEIDLPTTISLSDGEEPPYQGPCTL 120  
QY 133 QLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIGATCYGSGRMEG 192  
Db 121 QLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIGATCYGSGRMEG 180  
QY 193 PPPTYSEVIGHYGPSSFOHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKOKGHPL 252  
Db 181 PPPTYSEVIGHYGPSSFOHQSSGPPSLLEGLRLHHTHIAPLESAAIWSKEDKOKGHPL 240

RESULT 14  
US-10-390-045-12  
; Sequence 12, Application US/10390045  
; Publication No. US20030170713A1  
; GENERAL INFORMATION:  
; APPLICANT: SRIVASTAVA, SHIV  
; APPLICANT: XU, LINDA L.  
; APPLICANT: MOUL, JUDD W.  
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
; FILE REFERENCE: 04995.0057-00000  
; CURRENT APPLICATION NUMBER: US/10/390,045  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US/09/769,482  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,772

; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,045  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-390-045-12

Query Match 62.7%; Score 845.5; DB 12; Length 244;  
Best Local Similarity 67.6%; Pred. No. 1.6e-66;  
Matches 169; Conservative 21; Mismatches 53; Indels 7; Gaps 4;  
QY 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPSD 61  
DB 1 AELEFAQIIIIIVVMVVIVCLLNHYKVSFTRFNRPNQSRREDGLPQEGCLWPSD 60  
QY 62 STVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVQHEIDLPTISLSDGE 121  
DB 61 SAAPRLGASE--IMHAFSRDRFTAPSFQDRFRSFRQTPYVQHEIDLPTISLSDGE 118  
QY 122 EPPYQGPCTIQLRDPQEQLELNRESVRAPPNRTIFDSLDMSARL-GGPCPPSSNSGIS 180  
DB 119 EPPYQGPCTIQLRDPQEQLELNRESVRAPPNRTIFDSLDIDIAMYSGGPCPPSSNSGIS 178  
QY 181 ATCYGSGRMGEGPPPTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
DB 179 ASTCSSNGRMGEGPPPTYSEVNGHHPGASFLHHORS--NAHRGSRLOFQQ-NNAESTIVP 234  
QY 241 SKEKDQKQGH 250  
DB 235 IKGDKRKPGN 244

## RESULT 15

US-10-094-749-1836  
; Sequence 1836, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOYUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1836  
; LENGTH: 269  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-094-749-1836

Query Match 62.7%; Score 845.5; DB 12; Length 269;  
Best Local Similarity 67.6%; Pred. No. 1.8e-66;  
Matches 169; Conservative 21; Mismatches 53; Indels 7; Gaps 4;  
QY 2 AELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPSD 61  
DB 24 AELEFAQIIIIIVVMVVIVCLLNHYKVSFTRFNRPNQSRREDGLPQEGCLWPSD 83  
QY 62 STVSGNGIPEQVYAPPRDRLAVPPPAQRERFHRFQTPYVQHEIDLPTISLSDGE 121  
DB 84 SAAPRLGASE--IMHAFSRDRFTAPSFQDRFRSFRQTPYVQHEIDLPTISLSDGE 141  
QY 122 EPPYQGPCTIQLRDPQEQLELNRESVRAPPNRTIFDSLDMSARL-GGPCPPSSNSGIS 180  
DB 142 EPPYQGPCTIQLRDPQEQLELNRESVRAPPNRTIFDSLDIDIAMYSGGPCPPSSNSGIS 201  
QY 181 ATCYGSGRMGEGPPPTYSEVIGHYFGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240  
DB 202 ASTCSSNGRMGEGPPPTYSEVNGHHPGASFLHHORS--NAHRGSRLOFQQ-NNAESTIVP 257  
QY 241 SKEKDQKQGH 250  
DB 258 IKGDKRKPGN 267

Search completed: December 4, 2003, 19:53:23

Job time : 32 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 04:18:58 ; Search time 387 Seconds  
(without alignments)  
9111.986 Million cell updates/sec

Title: US-09-857-826B-44

Perfect score: 1061

Sequence: 1 tctctcttggttggtgta.....cttgcaaaaaaaaaaaaaa 1061

Scoring table:

OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2201672 seqs, 1661799599 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	90.6	969	11	US-09-796-753-55
2	926	87.3	1141	13	Sequence 55, Appl
3	926	87.3	1141	15	Sequence 208, App
4	926	87.3	1850	13	Sequence 412, App
5	926	87.3	4527	13	Sequence 44, Appl
6	912	86.0	1140	13	Sequence 2, Appli
7	902	85.0	4839	13	Sequence 1, Appli
8	902	85.0	4839	13	Sequence 119, App
9	902	85.0	4839	13	Sequence 84, Appl
10	891	84.0	1066	14	Sequence 85, Appl
11	788	74.3	806	13	Sequence 71, Appl
12	749	70.6	1321	10	Sequence 45, Appl
13	708	66.7	759	13	Sequence 1, Appli
14	701	66.1	861	10	Sequence 2, Appli
15	509	48.0	1583	15	Sequence 3, Appli
					Sequence 32, Appl

c	16	313	29.5	693	10	US-09-934-249-14	Sequence 14, Appl
	17	269	25.4	467	11	US-09-918-995-2074	Sequence 2074, Ap
	18	102	9.6	368	10	US-09-783-590-3464	Sequence 3464, Ap
	19	60	5.7	60	13	US-09-908-975-13620	Sequence 13620, A
	20	50	4.7	65	10	US-09-783-590-3488	Sequence 3488, Ap
	21	44	4.1	878	10	US-09-934-249-12	Sequence 12, Appl
	22	44	4.1	1713	11	US-09-796-753-57	Sequence 57, Appl
	23	32	3.0	577	9	US-09-864-761-20542	Sequence 20542, A
	24	32	3.0	1964	9	US-09-864-761-3776	Sequence 3776, Ap
	25	29	2.7	401	9	US-09-864-761-3936	Sequence 3936, Ap
	26	29	2.7	446	9	US-09-864-761-20699	Sequence 20699, A
	27	29	2.7	475	10	US-09-934-249-15	Sequence 15, Appl
c	28	29	2.7	3444	13	US-10-293-382-16	Sequence 16, Appl
c	29	26	2.5	1493	13	US-10-029-386-25133	Sequence 25133, A
c	30	26	2.5	7733	10	US-09-860-670-159	Sequence 159, App
	31	26	2.5	8429	9	US-09-738-885-3	Sequence 3, Appli
c	32	25	2.4	331	9	US-09-864-761-17053	Sequence 17053, A
c	33	25	2.4	465	9	US-09-864-761-230	Sequence 230, App
	34	25	2.4	470	9	US-09-864-761-3121	Sequence 3121, App
	35	25	2.4	522	9	US-09-864-761-19900	Sequence 19900, A
	36	25	2.4	522	13	US-10-029-386-22252	Sequence 22252, A
	37	25	2.4	571	13	US-10-029-386-9045	Sequence 9045, Ap
	38	25	2.4	756	13	US-10-029-386-20432	Sequence 20432, A
	39	24	2.3	24	13	US-10-390-045-5	Sequence 5, Appli
c	40	24	2.3	24	13	US-10-390-045-6	Sequence 6, Appli
c	41	24	2.3	24	13	US-10-390-045-10	Sequence 10, Appl
	42	24	2.3	58	15	US-10-106-698-3514	Sequence 3514, Ap
c	43	24	2.3	200	14	US-10-005-169-6	Sequence 6, Appli
c	44	24	2.3	569	13	US-10-029-386-3710	Sequence 3710, Ap
c	45	24	2.3	1884	9	US-09-853-386-62	Sequence 62, Appl

## ALIGNMENTS

RESULT 1  
US-09-796-753-55  
; Sequence 55, Application US/09796753  
; Publication No. US20030027998A1  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
; FILE REFERENCE: 7853-227-999  
; CURRENT APPLICATION NUMBER: US/09/796,753  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/183,175  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 09/223,094  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/223,546  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/224,246  
; PRIOR FILING DATE: 1998-12-30  
; PRIOR APPLICATION NUMBER: 09/259,388  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: 60/122,458  
; PRIOR FILING DATE: 1999-03-01  
; PRIOR APPLICATION NUMBER: 09/312,359  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 09/336,536  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: 09/342,687  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 09/345,464  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: 09/365,164  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: 09/399,723  
; PRIOR FILING DATE: 1999-09-20  
; PRIOR APPLICATION NUMBER: 09/409,634  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: 09/471,179  
; PRIOR FILING DATE: 1999-12-23

;; PRIOR APPLICATION NUMBER: 09/474,071  
;; PRIOR FILING DATE: 1999-12-29  
;; PRIOR APPLICATION NUMBER: 09/474,072  
;; PRIOR FILING DATE: 1999-12-29  
;; PRIOR APPLICATION NUMBER: 09/514,010  
;; PRIOR FILING DATE: 2000-02-25  
;; PRIOR APPLICATION NUMBER: 09/516,745  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 09/572,002  
;; PRIOR FILING DATE: 2000-05-14  
;; PRIOR APPLICATION NUMBER: 09/597,993  
;; PRIOR FILING DATE: 2000-06-19  
;; PRIOR APPLICATION NUMBER: 09/599,596  
;; PRIOR FILING DATE: 2000-06-22  
;; PRIOR APPLICATION NUMBER: 09/630,334  
;; PRIOR FILING DATE: 2000-07-31  
;; PRIOR APPLICATION NUMBER: 09/606,565  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/606,317  
;; PRIOR FILING DATE: 2000-06-29  
;; PRIOR APPLICATION NUMBER: 09/665,666  
;; PRIOR FILING DATE: 2000-09-20  
;; PRIOR APPLICATION NUMBER: 09/677,751  
;; PRIOR FILING DATE: 2000-09-30  
;; NUMBER OF SEQ ID NOS: 162  
;; SEQ ID NO 55  
;; LENGTH: 969  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (6)...(761)  
US-09-796-753-55

Query Match 90.6%; Score 961; DB 11; Length 969;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 961; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 101 ATGGGGAGCTGAGTTTCTCAGATCATCATCTGCTGGTGTATGATGTGTG 160  
DB 6 ATGGGGAGCTGAGTTTCTCAGATCATCATCTGCTGGTGTATGATGTGTG 65  
QY 161 GTGGTATCATCTGCTGCTGAGCCACTACAGCTGTCTGACGGTCTTTCATCAGCCGG 220  
DB 66 GTGGTATCATCTGCTGCTGAGCCACTACAGCTGTCTGACGGTCTTTCATCAGCCGG 125  
QY 221 CACAGCCAGGGCGGAGGAGAGATGCTCTCTCAGAGAGGATGCTGTGGCCCTCG 280  
DB 126 CACAGCCAGGGCGGAGGAGAGATGCTCTCTCAGAGAGGATGCTGTGGCCCTCG 185  
QY 281 GAGAGCACAGTGTACGGCAACGGATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 340  
DB 186 GAGAGCACAGTGTACGGCAACGGATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 245  
QY 341 ACCGACCGCTGCGCGCTGCGCCCTTCGCCCGAGGGGCTTCCACCGCTTCCAGCCC 400  
DB 246 ACCGACCGCTGCGCGCTGCGCCCTTCGCCCGAGGGAGGCTTCCACCGCTTCCAGCCC 305  
QY 401 ACCTATCCGTACTCTGAGCAGCAGATCGACCTGCGCCGCCACCATCTCTGCTGTGAGCGGG 460  
DB 306 ACCTATCCGTACTCTGAGCAGCAGATCGACCTGCGCCGCCACCATCTCTGCTGTGAGCGGG 365  
QY 461 GAGGAGCCCGCCCTTACAGGGCCCTGACCTCTCCAGTTCGGGACCCCGAGCAGCAG 520  
DB 366 GAGGAGCCCGCCCTTACAGGGCCCTGACCTCTCCAGTTCGGGACCCCGAGCAGCAG 425  
QY 521 CTGGAACTGAACCGGGAGTCTGGTGGCGGACCCCGAGCAGCAGTCTTTCAGCAGTAC 580  
DB 426 CTGGAACTGAACCGGGAGTCTGGTGGCGGACCCCGAGCAGCAGTCTTTCAGCAGTAC 485  
QY 581 CTGATGATGATGTCAGGCTGGGGCGGCCCTGCGCCCGCCAGCAGTAACTCGGGCATCAGC 640  
DB 486 CTGATGATGATGTCAGGCTGGGGCGGCCCTGCGCCCGCCAGCAGTAACTCGGGCATCAGC 545

QY 641 GCCACGTCTACGGCGGGCGCATGAGGGCGCGCCACCTACAGCGAGTCT 700  
DB 546 GCCACGTCTACGGCGGGCGCATGAGGGCGCGCCACCTACAGCGAGTCT 605  
QY 701 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGCAGTGGGCCCTCTCTTG 760  
DB 606 ATCGGCCACTACCGGGGTCTCTTCCAGCACAGCAGCAGTGGGCCCTCTCTTG 665  
QY 761 CTGAGGGGACCGGGCTCCACACACACATCGGGCCCTTAGAGAGCGCAGCATCTGG 820  
DB 666 CTGAGGGGACCGGGCTCCACACACACATCGGGCCCTTAGAGAGCGCAGCATCTGG 725  
QY 821 AGCAAGAGAGAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGCGGGCT 880  
DB 726 AGCAAGAGAGAGGATAAACAAGAGGACACCTCTCTAGGGTCCCGAGGGCGGGCT 785  
QY 881 GGGGCTGCTAGGTGAAAGGAGAGACACTCCGGCTTTCTTAGAGAGAGTGTAGAGGAA 940  
DB 786 GGGGCTGCTAGGTGAAAGGAGAGACACTCCGGCTTTCTTAGAGAGAGTGTAGAGGAA 845  
QY 941 GCGGGGGGCGAGCAACGATCTGTGGCGCTCCCTCCACCTCTCTGTGTATAATA 1000  
DB 846 GCGGGGGGCGAGCAACGATCTGTGGCGCTCCCTCCACCTCTCTGTGTATAATA 905  
QY 1001 TTTACATGTGATGTCTGTGCTGAATGCAAGAGCTTAAGAGAGCTTGCACAAAAA 1060  
DB 906 TTTACATGTGATGTCTGTGCTGAATGCAAGAGCTTGCACAAAAA 965  
QY 1061 A 1061  
DB 966 A 966

## RESULT 2

US-10-301-822-208  
; Sequence 208, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgett, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MP01-029P2RNM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 208  
; LENGTH: 1141  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (96)...(854)  
US-10-301-822-208

Query Match 87.3%; Score 926; DB 13; Length 1141;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 81 GTCTCTCGGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCTCGTG 140
Db 76 GTCTCTCGGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCTCGTG 135
QY 141 TGGTGATGATGCTGATGTTGGTGGTGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTG 200
Db 136 TGGTGATGATGCTGATGTTGGTGGTGATCAGCTGCTGCTGAGCCACTACAGAGCTGCTG 195
QY 201 CACGGTCTTTCATCAGCCGCGCACAGCAGGGCGGAGAGAGAAAGATGCCCTGCTCAG 260
Db 196 CACGGTCTTTCATCAGCCGCGCACAGCAGGGCGGAGAGAGAAAGATGCCCTGCTCAG 255
QY 261 AAGGATGCTGTGGCCCTCGAGAGACAGAGTGTTCAGGCAACGGAAATCCAGAGCCGAGG 320
Db 256 AAGGATGCTGTGGCCCTCGAGAGACAGAGTGTTCAGGCAACGGAAATCCAGAGCCGAGG 315
QY 321 TCTAGCCCGCGCTCGGCCACCGAGCTGGCGTGGCGTGGCGCTTCCGCCAGCGGAGC 380
Db 316 TCTAGCCCGCGCTCGGCCACCGAGCTGGCGTGGCGTGGCGCTTCCGCCAGCGGAGC 375
QY 381 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGCACGAGATCGACCTGCGGCCCA 440
Db 376 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGCACGAGATCGACCTGCGGCCCA 435
QY 441 CCATCTCGCTGTACAGCGGGAGAGCCGCCACCTTACCGAGGGCCCTTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTACAGCGGGAGAGCCGCCACCTTACCGAGGGCCCTTGCACCTCCAGC 495
QY 501 TTCGGGACCCCGAGCAGCTGGAACTGAACGGGAGTCGGTGGCGGCGGCGGCGGCGG 560
Db 496 TTCGGGACCCCGAGCAGCTGGAACTGAACGGGAGTCGGTGGCGGCGGCGGCGGCGG 555
QY 561 GAACCATCTTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 620
Db 556 GAACCATCTTTCAGAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 615
QY 621 GCAGTAACCTCGGCAATCAGGCGCACGCTGTACGCGAGCGGCGGCGGCGGCGGCGG 680
Db 616 GCAGTAACCTCGGCAATCAGGCGCACGCTGTACGCGAGCGGCGGCGGCGGCGGCGG 675
QY 681 CGCCCACTTACAGCGAGTCTATCGGCCACTACCGGGTCTCTTCCAGAGCAGCAGCA 740
Db 676 CGCCCACTTACAGCGAGTCTATCGGCCACTACCGGGTCTCTTCCAGAGCAGCAGCA 735
QY 741 CGAGTGGGCGGCGCTCTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 800
Db 736 CGAGTGGGCGGCGCTCTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 795
QY 801 TAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCCCTCTTAGG 860
Db 796 TAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCCCTCTTAGG 855
QY 861 GTCCCGAGGGGCGGCGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 920
Db 856 GTCCCGAGGGGCGGCGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 915
QY 921 TAGAAGAGAGTGTAGAGGAGGCGGGGCGGAGCAAGCATCTGTGTGGCCCTTCCCTCC 980
Db 916 TAGAAGAGAGTGTAGAGGAGGCGGGGCGGAGCAAGCATCTGTGTGGCCCTTCCCTCC 975
QY 981 CACCTCCCTGTGTATAAATATTTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1040
Db 976 CACCTCCCTGTGTATAAATATTTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1035
QY 1041 GCTTGCAGAAAAA 1057
Db 1036 GCTTGCAGAAAAA 1052
```

RESULT 3

US-10-205-823-412

; Sequence 412, Application US/10205823

```
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-205-823-412
```

```
Query Match 87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 81 GTCTCTCGGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCTCGTG 140
Db 76 GTCTCTCGGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCTCGTG 135
QY 141 TGGTGATGATGCTGATGTTGGTGGTGATCAGTGTCTGCTGAGCCACTACAGAGCTGCTG 200
Db 136 TGGTGATGATGCTGATGTTGGTGGTGATCAGTGTCTGCTGAGCCACTACAGAGCTGCTG 195
QY 201 CACGGTCTTTCATCAGCCGCGCACAGCAGGGCGGAGAGAGAAAGATGCCCTGCTCAG 260
Db 196 CACGGTCTTTCATCAGCCGCGCACAGCAGGGCGGAGAGAGAAAGATGCCCTGCTCAG 255
QY 261 AAGGATGCTGTGGCCCTCGAGAGACAGTGTTCAGGCAACGGAAATCCAGAGCCGAGG 320
Db 256 AAGGATGCTGTGGCCCTCGAGAGACAGTGTTCAGGCAACGGAAATCCAGAGCCGAGG 315
QY 321 TCTAGCCCGCGCTCGGCCACCGAGCTGGCGTGGCGTGGCGCTTCCGCCAGCGGAGC 380
Db 316 TCTAGCCCGCGCTCGGCCACCGAGCTGGCGTGGCGTGGCGCTTCCGCCAGCGGAGC 375
QY 381 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGCACGAGATCGACCTGCGGCCCA 440
Db 376 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGCACGAGATCGACCTGCGGCCCA 435
QY 441 CCATCTCGCTGTACAGCGGGAGAGCCGCCACCTTACCGAGGGCCCTTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTACAGCGGGAGAGCCGCCACCTTACCGAGGGCCCTTGCACCTCCAGC 495
QY 501 TTCGGGACCCCGAGCAGCTGGAACTGAACGGGAGTCGGTGGCGGCGGCGGCGGCGG 560
Db 496 TTCGGGACCCCGAGCAGCTGGAACTGAACGGGAGTCGGTGGCGGCGGCGGCGGCGG 555
```

Qy	561	GAACCATCTTCCGACAGTGAACCTGATGATAGTGCACGGCTGGGGGGCCCTCTGCCCCCCCA	620
Db	556	GAACCATCTTCCGACAGTGAACCTGATGATAGTGCACGGCTGGGGGGCCCTCTGCCCCCCCA	615
Qy	621	GCAGTAACTCGGGCATCAGCGCCACGTCCTACGGCAGCGGGGGCGCATGGAGGGGGCCGC	680
Db	616	GCAGTAACTCGGGCATCAGCGCCACGTCCTACGGCAGCGGGGGCGCATGGAGGGGGCCGC	675
Qy	681	CGCCCACTTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTTCCAGCACCCAGCAGA	740
Db	676	CGCCCACTTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTTCCAGCACCCAGCAGA	735
Qy	741	GCAGTGGGCGCCCTCTCTTCTGGAGGGGACCCGGCTCCACCAACACACATCGCGCCCC	800
Db	736	GCAGTGGGCGCCCTCTCTTCTGGAGGGGACCCGGCTCCACCAACACACATCGCGCCCC	795
Qy	801	TAGAGAGCGCAGCCATCTGGAGCAGAAAGATTAACAGAAAGGACACCTCTCTTAGG	860
Db	796	TAGAGAGCGCAGCCATCTGGAGCAGAAAGATTAACAGAAAGGACACCTCTCTTAGG	855
Qy	861	GTCCCCAGGGGGCGCGGCTCGGTAGGTGAAAGGACAGAACACATCCGCGCTTCT	920
Db	856	GTCCCCAGGGGGCGCGGCTCGGTAGGTGAAAGGACAGAACACATCCGCGCTTCT	915
Qy	921	TAGAAGAGGAGTGAGAGAGAAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCCCTCC	980
Db	916	TAGAAGAGGAGTGAGAGAGAAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCCCTCC	975
Qy	981	CACCTCCCTGTGTATAAATATTTACATGTGATGTCTGATGTCTGATGCACAAGCTAAGAGA	1040
Db	976	CACCTCCCTGTGTATAAATATTTACATGTGATGTCTGATGCACAAGCTAAGAGA	1035
Qy	1041	GCTTGCAAAAAA 1057	
Db	1036	GCTTGCAAAAAA 1052	
RESULT 4			
; Sequence 44, Application US/10241220			
; Publication No. US20030148408A1			
; GENERAL INFORMATION:			
; APPLICANT: Frantz, Gretchen			
; APPLICANT: Hillan, Kenneth J.			
; APPLICANT: Phillips, Heidi			
; APPLICANT: Polakis, Paul			
; APPLICANT: Spencer, Susan			
; APPLICANT: Williams, P. Mickey			
; APPLICANT: Wu, Thomas			
; APPLICANT: Zhang, Zemin			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF TUMOR			
; FILE REFERENCE: P5010R1-US			
; CURRENT APPLICATION NUMBER: US/10/241.220			
; CURRENT FILING DATE: 2002-12-13			
; NUMBER OF SEQ ID NOS: 120			
; SEQ ID NO 44			
; LENGTH: 1850			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
US-10-241-220-44			

Qy	201	CACGGTCCTTTTCATCAGCCGGGCA	CAGCCAGGGGCGGAGAGAAATGCCCTGTCTCTCAG	250
Db	196	CACGGTCCTTTTCATCAGCCGGGCA	CAGCCAGGGGCGGAGAGAAATGCCCTGTCTCTCAG	255
Qy	261	AAGGATGCTGTGCGCCTTCGGAGAGACACAGTGT	CAGGCAACGGAATCCAGAGCCGCGAGG	320
Db	256	AAGGATGCTGTGTGGCCTTCGGAGAGACACAGTGT	CAGGCAACGGAATCCAGAGCCGCGAGG	315
Qy	321	TCTACGCCCGCCTTCGGCCCA	CAGCGCTGTGCGCCCTTCGCCCAGCGGAGC	380
Db	316	TCTACGCCCGCCTTCGGCCCA	CAGCGCTGTGCGCCCTTCGCCCAGCGGAGC	375
Qy	381	GCTTCCACCGCTTTCAGACCCACCTATCTCGTAC	CTGACGACGAGATCGACCTGGCGCCCA	440
Db	376	GCTTCCACCGCTTTCAGACCCACCTATCTCGTAC	CTGACGACGAGATCGACCTGGCACCCA	435
Qy	441	CCATCTCGCTGTTCAGACGGGGAGAGCCCCAC	CCCTACAGGGCGCCCTGTCACCTCCAGC	500
Db	436	CCATCTCGCTGTTCAGACGGGGAGAGCCCCAC	CCCTACAGGGCGCCCTGTCACCTCCAGC	495
Qy	501	TTCCGGACCCCGAGCAGCTGGAACTGAA	CCGGAGTCTGTGGGGACCCCCCAACA	560
Db	496	TTCCGGACCCCGAGCAGCTGGAACTGAA	CCGGAGTCTGTGGCGCACCCCAACA	555
Qy	561	GAACCATCTTCGACAGTGA	CTCATGTAGTCCAGGCTGGCGGGCCCTGCCCCC	620
Db	556	GAACCATCTTCGACAGTGA	CTCATGTAGTCCAGGCTGGCGGGCCCTGCCCCC	615
Qy	621	GCAGTAACCTCGGGCATCAGCCCA	CTGCTGTAACGGCAGCGCGGGCGCATGGAGGGCGCG	680
Db	616	GCAGTAACCTCGGGCATCAGCCCA	CTGCTGTAACGGCAGCGCGGGCGCATGGAGGGCGCG	675
Qy	681	CGCCCACTTACAGGAGTTCATCGGCCAC	TACCCGGGTCTCTCTTCCAGCACACGACAGA	740
Db	676	CGCCCACTTACAGGAGTTCATCGGCCAC	TACCCGGGTCTCTCTTCCAGCACACGACAGA	735
Qy	741	GCAGTGGCGCGCCTCTCTTCTGGAGGGGAC	CCCGGTCCACACACATCGCGCCCC	800
Db	736	GCAGTGGCGCGCCTCTCTTCTGGAGGGGAC	CCCGGTCCACACACATCGCGCCCC	795
Qy	801	TAGAGAGCGAGCGCATCTGGAGCA	AAAGAGATAAACGAAAGGACACCTCTCTAGG	860
Db	796	TAGAGAGCGAGCGCATCTGGAGCA	AAAGAGATAAACGAAAGGACACCTCTCTAGG	855
Qy	861	GTCCCCAGGGGGCGCGGCTGGGGCT	TGGGCTGCGTAGTGAAAGGACACATCCCGGTTCT	920
Db	856	GTCCCCAGGGGGCGCGGCTGGGGCT	TGGGCTGCGTAGTGAAAGGACACATCCCGGTTCT	915
Qy	921	TAGAAGAGGAGTGAGAGGA	AGCGGGGGCGCAGCAACGCGATCGTGTGGCCCTCCCGCTCC	980
Db	916	TAGAAGAGGAGTGAGAGGA	AGCGGGGGCGCAGCAACGCGATCGTGTGGGCCCTCCCGCTCC	975
Qy	981	CACCTCCCTGTGTATAATATTTTACATGT	GATGTCTGTGTGATGTCGAAAGGACAACTGAACTAAGAGA	1040
Db	976	CACCTCCCTGTGTATAATATTTTACATGT	GATGTCTGTGTGATGTCGAAAGGACAACTGAACTAAGAGA	1035
Qy	1041	GCTTGCACAAAAA	AAAAA 1057	
Db	1036	GCTTGCACAAAAA	AAAAA 1052	



NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 4527  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (96)...(851)  
US-09-821-812-2

Query Match 87.3%; Score 926; DB 13; Length 4527;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 81 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 140  
DB 76 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 135  
QY 141 TGGTGATGATGTGATGTGTGTGTGATCATCGTCTGTGAGCACTACAAAGCTGTGTG 200  
DB 136 TGGTGATGATGTGATGTGTGTGTGATCATCGTCTGTGAGCACTACAAAGCTGTGTG 195  
QY 201 CACGGTCTTTCATCAGCGGCACAGCCAGGGCGGAGGAGAGAGAGTGCCTCTCTCAG 260  
DB 196 CACGGTCTTTCATCAGCGGCACAGCCAGGGCGGAGGAGAGAGAGTGCCTCTCTCAG 255  
QY 261 AAGGATGCTGTGGCCCTCGAGAGCACAGTGTCAAGGCAACGGAATCCAGAGCCGAGG 320  
DB 256 AAGGATGCTGTGGCCCTCGAGAGCACAGTGTCAAGGCAACGGAATCCAGAGCCGAGG 315  
QY 321 TCTACGCCCGCCTCGGCCACCGACCGCTGCGCCCTGCGCCCTTGCAGCGGAGC 380  
DB 316 TCTACGCCCGCCTCGGCCACCGACCGCTGCGCCCTTGCAGCGGAGC 375  
QY 381 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGACACAGATCGACCTGCCCA 440  
DB 376 GCTTCCACCGCTTCCAGCCCACTATCCGTACTGTCAGACACAGATCGACCTGCCCA 435  
QY 441 CCATCTCGTGTGACAGCGGGAGGAGCCCACTTACAGGGCCCTTGCAGCGGCTCCAGC 495  
DB 436 CCATCTCGTGTGACAGCGGGAGGAGCCCACTTACAGGGCCCTTGCAGCGGCTCCAGC 495  
QY 501 TTGCGGACCCGAGCAGCAGCTGGAATGAAACCGGAGTGTGTCGCGCACCCCAACA 560  
DB 496 TTGCGGACCCGAGCAGCAGCTGGAATGAAACCGGAGTGTGTCGCGCACCCCAACA 555  
QY 561 GAACCATCTTCACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 620  
DB 556 GAACCATCTTCACAGTACCTGATGATGATGATGATGATGATGATGATGATGATGATG 615  
QY 621 GCAGTAACTCGGGCATTCAGCGCCACGTGCTACGGCAGCGGGCGGCGATGAGAGGGCG 680  
DB 616 GCAGTAACTCGGGCATTCAGCGCCACGTGCTACGGCAGCGGGCGGCGATGAGAGGGCG 675  
QY 681 CGCCCACTTACAGCAGGCTCATCGGCACCTACCCGGGTCTCTTTCAGACACAGCAGA 740  
DB 676 CGCCCACTTACAGCAGGCTCATCGGCACCTACCCGGGTCTCTTTCAGACACAGCAGA 735  
QY 741 GCAGTGGGCGCCCTCTTGTGTGAGGGAGCCCGGCTCCACACACACATCGGCGCC 800  
DB 736 GCAGTGGGCGCCCTCTTGTGTGAGGGAGCCCGGCTCCACACACACATCGGCGCC 795  
QY 801 TAGAGAGCGACCCATCTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGG 860  
DB 796 TAGAGAGCGACCCATCTGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTCTAGG 855  
QY 861 GTCCCCAGGGGCGGGGTGCGGTAGGTGAAAGGAGGAGCACTCCGCGCTTCT 920  
DB 856 GTCCCCAGGGGCGGGGTGCGGTAGGTGAAAGGAGGAGCACTCCGCGCTTCT 915  
QY 921 TAGAAGAGAGTGTAGAGGAGCGGGGCGGAGCAACGCATGCTGTGCGCCCTCCCTCC 980

DB 916 TAGAAGAGAGTGTAGAGGAGCGGGGCGGAGCAACGCATCGTGTGGCCCTCCCTCC 975  
QY 981 CACCTCCCTGTGTATAAATATTTACATGTGATGTGTGTGTGTGTGTGTGTGTGTGTG 1040  
DB 976 CACCTCCCTGTGTATAAATATTTACATGTGATGTGTGTGTGTGTGTGTGTGTGTGTG 1035  
QY 1041 GCTTGCAAAAAA 1057  
DB 1036 GCTTGCAAAAAA 1052

RESULT 6  
US-10-390-045-1  
Sequence 1, Application US/10390045  
Publication No. US20030170713A1  
GENERAL INFORMATION:  
APPLICANT: SRIVASTAVA, SHIV  
APPLICANT: MOUL, JUDD W.  
APPLICANT: XU, LINDA L.  
APPLICANT: SEGAWA, TAKEHIKO  
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED  
TITLE OF INVENTION: POYNUCLEOTIDE ARRAY  
FILE REFERENCE: 04995.0057-00000  
CURRENT APPLICATION NUMBER: US/10/390,045  
PRIOR FILING DATE: 2003-03-18  
PRIOR APPLICATION NUMBER: US/09/169,482  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/178,772  
PRIOR FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/179,045  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1140  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (95)...(850)  
US-10-390-045-1

Query Match 86.0%; Score 912; DB 13; Length 1140;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 95 CAGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCAT 154  
DB 89 CAGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCAT 148  
QY 155 ATGGTGGTGTGATCAGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATC 214  
DB 149 ATGGTGGTGTGATCAGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATC 208  
QY 215 AGCCGGCA CAGCCAGGGCGGAGGAGAGAGTGCCTGTCTCTCAGAGGATGCCTGTGG 274  
DB 209 AGCCGGCACAGCCAGGGCGGAGGAGAGAGTGCCTGTCTCTCAGAGGATGCCTGTGG 268  
QY 275 CCCTCGGAGACACAGTGTCAAGCAACCGAATCCAGAGCCGAGCTACGCCCGGCT 334  
DB 269 CCCTCGGAGACACAGTGTCAAGCAACCGAATCCAGAGCCGAGCTACGCCCGGCT 328  
QY 335 CGGCCACCGAGCCGCTGCGCGCTTCCGCCAGCGGAGCGCTTCCACCGCTTC 394  
DB 329 CGGCCACCGAGCCGCTGCGCGCTTCCGCCAGCGGAGCGCTTCCACCGCTTC 388  
QY 395 CAGCCCACTATCCGTACTCTGACGACGAGATCGA CTTGCCGCCACCACCATCTCGTGTCA 454  
DB 389 CAGCCCACTATCCGTACTCTGACGACGAGATCGA CTTGCCGCCACCACCATCTCGTGTCA 448  
QY 455 GACGGGAGGAGCCCCCACCCTACAGGGCCCTTGCACCTCCAGCTTCGGGACCCCGAG 514  
DB 449 GACGGGAGGAGCCCCCACCCTACAGGGCCCTTGCACCTCCAGCTTCGGGACCCCGAG 508

QY 515 CAGACCTGGAAGTGAACCGGAGTCGGTGGCGGACCCCAACAGACCACTCTTCGAC 574  
Db 509 CAGCAGCTGGAAGTGAACCGGAGTCGGTGGCGGACCCCAACAGACCACTCTTCGAC 568  
QY 575 AGTGACCTGATGATAGTCCAGGCTGGGCGGCCCTGCCCCCGCCAGAGTAACCTCGGGC 634  
Db 569 AGTGACCTGATGATAGTCCAGGCTGGGCGGCCCTGCCCCCGCCAGAGTAACCTCGGGC 628  
QY 635 ATCAGCGCACGTGCTACGGCAGCGCGCGGCATGAGGGCGCCCGCCACCTACAGC 694  
Db 629 ATCAGCGCACGTGCTACGGCAGCGCGCGGCATGAGGGCGCCCGCCACCTACAGC 688  
QY 695 GAGGTCATCGGCACCTACCGGGGTCTCTTCCAGCACCAGAGAGTGGGCGGCC 754  
Db 689 GAGGTCATCGGCACCTACCGGGGTCTCTTCCAGCACCAGAGAGTGGGCGGCC 748  
QY 755 TCCCTTGCTGAGGGAGCCCGGCTCCACACACACACATCGCGCCCTAGAGAGCGCAGCC 814  
Db 749 TCCCTTGCTGAGGGAGCCCGGCTCCACACACACACATCGCGCCCTAGAGAGCGCAGCC 808  
QY 815 ATCTGAGCAAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGC 874  
Db 809 ATCTGAGCAAGAGAGATAAACAGAAAGACACCTCTCTAGGGTCCCGAGGGGGC 868  
QY 875 CGGGCTGGGGCTCGCTAGTGAAGGAGCAACACTCCCGGCTCTTAGAAGAGAGTGA 934  
Db 869 CGGGCTGGGGCTCGCTAGTGAAGGAGCAACACTCCCGGCTCTTAGAAGAGAGTGA 928  
QY 935 GAGGAAGCGGGGGCGGAGCAACGATCGTGTGGCCCTCCACCTCCCTCTGTGA 994  
Db 929 GAGGAAGCGGGGGCGGAGCAACGATCGTGTGGCCCTCCACCTCCCTCTGTGA 988  
QY 995 TAAATATTTACATGTGATGTCTGCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1054  
Db 989 TAAATATTTACATGTGATGTCTGCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1048  
QY 1055 AAA 1057  
Db 1049 AAA 1051

## RESULT 7

US-10-241-220-119  
; Sequence 119, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 119  
; LENGTH: 4839  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-241-220-119

Query Match 85.0%; Score 902; DB 13; Length 4839;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCTCGTGGTGGTGGTGGTGGTGGTGG 164

Db 430 CGGAGCTGGAGTTGTTTCAGATCATCATCTCGTGGTGGTGGTGGTGGTGGTGG 489  
QY 165 TGATCAGCTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTCTTCAACAGCCGACCA 224  
Db 490 TGATCAGCTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTCTTCAACAGCCGACCA 549  
QY 225 GCCAGGCGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGA 284  
Db 550 GCCAGGCGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCCCTCGGAGA 609  
QY 285 GCACAGTGTACAGCAACCGAAATCCAGAGCGCAGGTCTACGCCCGGCTCGGCCACCG 344  
Db 610 GCACAGTGTACAGCAACCGAAATCCAGAGCGCAGGTCTACGCCCGGCTCGGCCACCG 669  
QY 345 ACCGCTGCGCTGCGGCCCTTTCGCGCAGCGGAGCGCTTCAACCGCTTCAGGCCACCT 404  
Db 670 ACCGCTGCGCTGCGGCCCTTTCGCGCAGCGGAGCGCTTCAACCGCTTCAGGCCACCT 729  
QY 405 ATCCGTACTCTCAGCAGCAGATCGACCTGCGCGCCACCATCTCGCTGTACAGCGGGAGG 464  
Db 730 ATCCGTACTCTCAGCAGCAGATCGACCTGCGCGCCACCATCTCGCTGTACAGCGGGAGG 789  
QY 465 AGCCCCACCTTACAGGGGCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524  
Db 790 AGCCCCACCTTACAGGGGCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 849  
QY 525 AACTGAACCGGGAGTGGTGGCGGACCCCAACAGAAACCATCTTTCACAGTGCACCTGA 584  
Db 850 AACTGAACCGGGAGTGGTGGCGGACCCCAACAGAAACCATCTTTCACAGTGCACCTGA 909  
QY 585 TGGATAGTGCAGAGTGGGCGGCCCTGCGCGCCAGCAGTAACCTCGGGCATCAGCGCA 644  
Db 910 TGGATAGTGCAGAGTGGGCGGCCCTGCGCGCCAGCAGTAACCTCGGGCATCAGCGCA 969  
QY 645 CGTGTACGCGAGCGGGCGGCATGAGAGGGCGCGCATGAGAGGGCGCGCCCATCAGCTCG 704  
Db 970 CGTGTACGCGAGCGGGCGGCATGAGAGGGCGCGCATGAGAGGGCGCGCCCATCAGCTCG 1029  
QY 705 GCCACTACCGGGGTCTCTTTCAGACACAGCAGAGAGAGTGGCGGCCCTCTCTTGTCTGG 764  
Db 1030 GCCACTACCGGGGTCTCTTTCAGACACAGCAGAGAGAGTGGCGGCCCTCTCTTGTCTGG 1089  
QY 765 AGGGACCGCGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCATCTCGAGCA 824  
Db 1090 AGGGACCGCGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCATCTCGAGCA 1149  
QY 825 AAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGG 884  
Db 1150 AAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGG 1209  
QY 885 CTGCGTAGGTGAAAGGAGAGAACTCCGCGCTTCTTAGAAGAGAGTGAAGAGAGCGG 944  
Db 1210 CTGCGTAGGTGAAAGGAGAGAACTCCGCGCTTCTTAGAAGAGAGTGAAGAGAGCGG 1269  
QY 945 GGGGCGCAGCAACGCGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTTA 1004  
Db 1270 GGGGCGCAGCAACGCGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTTA 1329  
QY 1005 CATGTGATGTCTGTGCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1057  
Db 1330 CATGTGATGTCTGTGCTGAATGCAAGCTAAGAGAGCTTGCAAAAAA 1382

## RESULT 8

US-10-269-909-84  
; Sequence 84, Application US/10269909  
; Publication No. US20030180747A1  
; GENERAL INFORMATION:  
; APPLICANT: HRUBAN, RALPH H.  
; APPLICANT: ARGANI, PEDRAM  
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE  
; APPLICANT: MAITRA, ANIRBAN  
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES

[illegible]

		TGGATAGTGCAGGCTGGGGGGCCCTTGCGCCCCCAGCAGTAACCTGGGTCATCAGGGCCA	969
Db			
		CGTGCTACGCACGCGGGCGGCATGAGGGGCCGCCACCCTACAGCGAGGTCATCG	704
Qy			
		CGTGCTACGGACGCGGGCGGCATGAGGGGCCGCCACCCTACAGCGAGGTCATCG	1029
Db			
		GCCACTACCCGGGGTCCTCTTCAGACACAGACAGCAGTGGGCGGCCCTCTCTGCTGG	764
Qy			
		GCCACTACCCGGGGTCCTCTTCAGACACAGACAGCAGTGGGCGGCCCTCTCTGCTGG	1089
Db			
		AGGGGACCCGGCTCCACCAACACATCGCGCCCTTAGAGAGCGCAGCCATCTGAGACA	824
Qy			
		AGGGGACCCGGCTCCACCAACACATCGCGCCCTTAGAGAGCGCAGCCATCTGAGACA	1149
Db			
		AAGNAGGATAAACAGAAAGCACCCCTCTTAGGGTCCCACAGGGGGCGCGGCTGGGG	884
Qy			
		AAGNAGGATAAACAGAAAGCACCCCTCTTAGGGTCCCACAGGGGGCGCGGCTGGGG	1209
Db			
		CTCGTAGGTGAAGAAGCAGAACACTCCGCGCTTTCTTAGAAGAGGAGCTGAGAGGAGGCG	944
Qy			
		CTCGTAGGTGAAGAAGCAGAACACTCCGCGCTTTCTTAGAAGAGGAGCTGAGAGGAGGCG	1269
Db			
		GGGGGCGCAGCAACGCTCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTTA	1004
Qy			
		GGGGGCGCAGCAACGCTCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTTA	1329
Db			
		CATGTGATGTGCTGAATGCAACAGCTTAAGAGAGCTTGCAAAAAA	1057
Qy			
		CATGTGATGTGCTGAATGCAACAGCTTAAGAGAGCTTGCAAAAAA	1382
Db			

**RESULT 10**

```

US-10-098-841-71
; Sequence 71, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 71
; LENGTH: 1066
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)..(867)

```

; Sequence 45, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; TREATMENT OF TUMOR  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 45  
; LENGTH: 806  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-241-220-45

Query Match 74.3%; Score 788; DB 13; Length 806;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 788; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GTCTCTCGAAACCCAGGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 140  
DB 13 GTCTCTCGAAACCCAGGCAATGGCGAGCTGGAGTTGTTTCAGATCATCATCATCGTGG 72

QY 141 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 200  
DB 73 TGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 132

QY 201 CACGGTCTTCATCAGCCGGGACAGCCAGGGGCGAGGAGAGAGAGAGAGAGAGAGAGAGAG 260  
DB 133 CACGGTCTTCATCAGCCGGGACAGCCAGGGGCGAGGAGAGAGAGAGAGAGAGAGAGAG 192

QY 261 AAGGATGCTGTGGCCCTCGGAGAGACAGATGTCAGGCAACGGAATCCAGAGCCGACAGG 320  
DB 193 AAGGATGCTGTGGCCCTCGGAGAGACAGATGTCAGGCAACGGAATCCAGAGCCGACAGG 252

QY 321 TCTACGCCCCCGCTCGGCCCCACCGACCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 380  
DB 253 TCTACGCCCCCGCTCGGCCCCACCGACCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGG 312

QY 381 GCTTCACGGCTTCAGCCGACCTATCGTACCTGACGACAGATCGACCTGCGGCCCA 440  
DB 313 GCTTCACGGCTTCAGCCGACCTATCGTACCTGACGACAGATCGACCTGCGGCCCA 372

QY 441 CCATCTCGTGTGACAGCGGGAGGAGCCGCCCTACCGAGGCGCCCTGACCCCTCCAGC 500  
DB 373 CCATCTCGTGTGACAGCGGGAGGAGCCGCCCTACCGAGGCGCCCTGACCCCTCCAGC 432

QY 501 TTCCGGGACCCCGAGCAGCAGCTGGAATGAACTGAACTGAACTGAACTGAACTGAACTGAA 560  
DB 433 TTCCGGGACCCCGAGCAGCAGCTGGAATGAACTGAACTGAACTGAACTGAACTGAACTGAA 492

QY 561 GAAACCATTTTCAGACAGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 620  
DB 493 GAAACCATTTTCAGACAGTGAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 552

QY 621 GAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGCGCGCGCATGAGAGGGGCGCG 680  
DB 553 GAGTAACCTCGGGCATCAGCGCCACCTGCTACGCGAGCGCGCGCGCATGAGAGGGGCGCG 612

QY 681 CGCCCACTTACAGCAGGATCATCGGCCACTACCCGGGGTCTCTCTTTCAGACACCGACAGA 740  
DB 613 CGCCCACTTACAGCAGGATCATCGGCCACTACCCGGGGTCTCTCTTTCAGACACCGACAGA 672

QY 741 GCAGTGGGCGCCCTCTCTTGTGGAGGGAGCCCGGCTCCACACACATCGCGCCCG 800

DB 673 CGAGTGGGCCGCCCTCTTGTCTGGAGGGAGCCCGCTCCACACACACATCGCGCCCC 732  
QY 801 TAGAGAGCGCAGCCATCTGGAGCAAGAGAGATATAACAGAAAGGACACCCCTCTCTAGG 860  
DB 733 TAGAGAGCGCAGCCATCTGGAGCAAGAGAGATATAACAGAAAGGACACCCCTCTCTAGG 792  
QY 861 GTCCCCCAG 868  
DB 793 GTCCCCCAG 800

## RESULT 12

US-09-934-249-1  
; Sequence 1, Application US/09934249  
; Patent No. US20020115081A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Richard T.  
; APPLICANT: Landschulz, Katherine T.  
; APPLICANT: Turi, Thomas G.  
; APPLICANT: Thompson, John F.  
; APPLICANT: Kennedy, Scott P.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS  
; FILE REFERENCE: P0738/7001/ERP/KA  
; CURRENT APPLICATION NUMBER: US/09/934,249  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/227,159  
; PRIOR FILING DATE: 2000-08-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1321  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (413)...(1273)  
US-09-934-249-1

Query Match 70.6%; Score 749; DB 10; Length 1321;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 799; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCATCATCATCAT 164  
DB 522 CGGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCATCATCATCAT 581

QY 165 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGACCA 224  
DB 582 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCGACCA 641

QY 225 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGAGGATGCTGTGGCCCTCGGAGA 284  
DB 642 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAGAGGATGCTGTGGCCCTCGGAGA 701

QY 285 GCACAGTGTCTAGGCAACGGAATCCAGAGCCGACAGTCTACGCCCGCCCTCGGCCCCACCG 344  
DB 702 GCACAGTGTCTAGGCAACGGAATCCAGAGCCGACAGTCTACGCCCGCCCTCGGCCCCACCG 761

QY 345 ACCGCTTGGCGGTGCGGCCCTTCGCCCHAGGGGAGCGCTTCACCGCTTCAGAGCCACCT 404  
DB 762 ACCGCTTGGCGGTGCGGCCCTTCGCCCHAGGGGAGCGCTTCACCGCTTCAGAGCCACCT 821

QY 405 ATCCGTACTCTGAGCAGCAGAGATCGACCTGCGGCCACCATCTCGCTGTCTCAGACGGGAGG 464  
DB 822 ATCCGTACTCTGAGCAGCAGAGATCGACCTGCGGCCACCATCTCGCTGTCTCAGACGGGAGG 881

QY 465 AGCCCCCACCCTTACAGGGGCCCTTCAGCCCTTCAGCTTCGGGACCCCGAGCAGCAGCTGG 524  
DB 882 AGCCCCCACCCTTACAGGGGCCCTTCAGCCCTTCAGCTTCGGGACCCCGAGCAGCAGCTGG 941

QY 525 AACTGAAACGGGAGTGGTGGCGGACCCCGCCAAACAGAAACCATCTTCAGACAGTGCCTGA 584



```
Db 110 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGTGTGATGATGTGATGTTGGTGG 169
Qy 165 TGATCAGCTGCTGTGAGCCACTAAAGCTGTCTGACAGTCTTTCATCAGCCGCGACA 224
Db 170 TGATCAGCTGCTGTGAGCCACTAAAGCTGTCTGACAGTCTTTCATCAGCCGCGACA 229
Qy 225 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGGAGA 284
Db 230 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAGGATGCTGTGGCCCTCGGAGA 289
Qy 285 GCACAGTGTGAGGCAACGAATCCAGAGCCGACAGTCTACGCCCGCTCGGCCACCG 344
Db 290 GCACAGTGTGAGGCAACGAATCCAGAGCCGACAGTCTACGCCCGCTCGGCCACCG 349
Qy 345 ACCGCTGCGCTGCGCCCTTCGCGCAGCGGAGCGCTTCCACGCTTCCAGCCCACT 404
Db 350 ACCGCTGCGCTGCGCCCTTCGCGCAGCGGAGCGCTTCCACGCTTCCAGCCCACT 409
Qy 405 ATCCGTACTGACAGCAGAGATCGACCTGCGCGCCACCATCTCGCTGTGACAGCGGAGG 464
Db 410 ATCCGTACTGACAGCAGAGATCGACCTGCGCGCCACCATCTCGCTGTGACAGCGGAGG 469
Qy 465 AGCCCCACCTTACCAGAGCCCTGCGCCCTGCGCCCTGCGGAGCCCGGAGCGAGCTGG 524
Db 470 AGCCCCACCTTACCAGAGCCCTGCGCCCTGCGCCCTGCGGAGCCCGGAGCGAGCTGG 529
Qy 525 AACTGAACCGGAGTCTGCTGCGGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 584
Db 530 AACTGAACCGGAGTCTGCTGCGGACCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 589
Qy 585 TGGATAGTGCAGGCTGCGGCGCCCTGCGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 644
Db 590 TGGATAGTGCAGGCTGCGGCGCCCTGCGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 649
Qy 645 CGTGTCTACGGCAGCGGCGGCGCATGGAGGCGCGCGCCCGGAGGAGGAGGAGGAGGAGG 704
Db 650 CGTGTCTACGGCAGCGGCGGCGCATGGAGGCGCGCGCCCGGAGGAGGAGGAGGAGGAGG 709
Qy 705 GCCACTACCGGGTCTCTCTTCCAGCACACACATGCGCGCGCGCGCGCGCGCGCGCGCG 764
Db 710 GCCACTACCGGGTCTCTCTTCCAGCACACACATGCGCGCGCGCGCGCGCGCGCGCGCG 769
Qy 765 AGGGAGCCCGGCTCCACACACACACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 824
Db 770 AGGGAGCCCGGCTCCACACACACACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 829
Qy 825 AAGAGAGGATAAACAGAAAGGACACCTCTC 856
Db 830 AAGAGAGGATAAACAGAAAGGACACCTCTC 861
```

## RESULT 15

```
US-10-000-256A-32
; Sequence 32, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 1583
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-000-256A-32
```

```
Query Match 48.0%; Score 509; DB 15; Length 1583;
Best Local Similarity 99.8%; Pred. No. 6.3e-228;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 313 GCGCAGAGTCTACGCGCCGCTCGGCCCCACGACCGCTGCGCGCTGCGCGCTTCCGCCCA 372
Db 827 GCGCAGAGTCTACGCGCCGCTCGGCCCCACGACCGCTGCGCGCTGCGCGCTTCCGCCCA 886
Qy 373 GCGGAGAGCGCTTCCACCGCTTCCAGCCCGCTATCCGTACCTGACAGCAGAGATGCACCT 432
Db 887 GCGGAGAGCGCTTCCACCGCTTCCAGCCCGCTATCCGTACCTGACAGCAGAGATGCACCT 946
Qy 433 GCGGCCCAACATCTGCTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 492
Db 947 GCGGCCCAACATCTGCTGTGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1006
Qy 493 CCTCCAGTCTCGGAGCCCGGAGCAGCAGCTGGAACCTGAAACCGGAGTCTGTCGCGGACCC 552
Db 1007 CCTCCAGTCTCGGAGCCCGGAGCAGCAGCTGGAACCTGAAACCGGAGTCTGTCGCGGACCC 1066
Qy 553 CCCAAACAGAACCATCTTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTG 612
Db 1067 CCCAAACAGAACCATCTTTCGACAGTACCTGATGATAGTGCAGGCTGGGCGGCCCTG 1126
Qy 613 CCCCCCAGCAGTAACCTCGGCGCATCAGCGCCACCGTCTACGSCAGCGCGCGCATGGA 672
Db 1127 CCCCCCAGCAGTAACCTCGGCGCATCAGCGCCACCGTCTACGSCAGCGCGCGCATGGA 1186
Qy 673 GGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 732
Db 1187 GGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1246
Qy 733 CCAGCAGAGCAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 792
Db 1247 CCAGCAGAGCAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1306
Qy 793 CCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
Db 1307 CCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1366
Qy 853 TCTCTAGGGTCCCGAGGGGG 872
Db 1367 TCTCTAGGGTCCCGAGGGGG 1386
```

Search completed: December 8, 2003, 06:25:42  
Job time : 391 secs

**THIS PAGE BLANK (USPTO)**